



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification ⁵ : C12N 15/12, C07K 13/00 C12N 5/10, A61K 37/02 G01N 33/68</p>	<p>A2</p>	<p>(11) International Publication Number: WO 94/11502 (43) International Publication Date: 26 May 1994 (26.05.94)</p>
<p>(21) International Application Number: PCT/GB93/02367 (22) International Filing Date: 17 November 1993 (17.11.93) (30) Priority data: 9224057.1 17 November 1992 (17.11.92) GB 9304677.9 8 March 1993 (08.03.93) GB 9304680.3 8 March 1993 (08.03.93) GB 9311047.6 28 May 1993 (28.05.93) GB 9313763.6 2 July 1993 (02.07.93) GB 9316099.2 3 August 1993 (03.08.93) GB 9321344.5 15 October 1993 (15.10.93) GB</p> <p>(71) Applicant (for all designated States except US): LUDWIG INSTITUTE FOR CANCER RESEARCH [GB/GB]; St. Mary's Hospital Medical School, Norfolk Place, Paddington, London W2 1PG (GB).</p>		<p>(72) Inventors; and (75) Inventors/Applicants (for US only) : MIYAZONO, Kohei [JP/SE]; Flogstavägen 63D, S-752 63 Uppsala (SE). DIJKE, Peter, Ten [NL/SE]; Flogstavägen 25C, S-752 63 Uppsala (SE). FRANZEN, Petra [SE/SE]; Lindsbergsgatan 15b, S-752 40 Uppsala (SE). YAMASHITA, Hidetoshi [JP/SE]; Flogstavägen 33A, S-752 63 Uppsala (SE). HELDIN, Carl-Henrik [SE/SE]; Hesselmans väg 35, S-752 63 Uppsala (SE).</p> <p>(74) Agent: GILL JENNINGS & EVERY; Broadgate House, 7 Eldon Street, London EC2M 7LH (GB).</p> <p>(81) Designated States: AU, CA, JP, KR, NZ, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).</p> <p>Published Without international search report and to be republished upon receipt of that report.</p>

(54) Title: ACTIVIN RECEPTOR-LIKE KINASES, PROTEINS HAVING SERINE THREONINE KINASE DOMAINS AND THEIR USE

```

cons.aa      C G G V          A K          E
htGFBR-II   LDTLVGRGRFAEYVYKAKLKQNTSEQFETVAVKI FPYDHYASWDRKDI FSDINLGHENILQF
mActr-IIB   LLEIKARGRFQCVWKAQLMN-----DFVAVKI KPLQDKQSWQSEREI FSTPGMGHENILQF
mActr-II    LLEVKARGRFQCVWKAQLLN-----EYVAVKI FPIODKQSWQNEYEVYSI PGMGHENILQF
daf-1       LKRVGSGRFGVNSRQDYRG-----EAVAVKVFNAIDEPAPFKKEIEI FETRMLRHPHVLRY
subdomains  I                II          III          IV

htGFBR-II   LTAERKTELGRQYMLITAFHAKGNLQEYLTRHVISWEDLRNVGSSLARGLSHLHSDHTP-C
mActr-IIB   IAAEKRGSNLEVELMLITAFHDKGSLIDYLGKNIITWNELCHVAETMSRGI SYLHEDVPMCR
mActr-II    IGAEKRGTSVDVQLMLITAFHEKGSLSDFLKAHVVSWNELCHIAETMARGLAYLHEDI PGLK
daf-1       IGSRRVDVTFVTELMLVIEYHPGSLHDFLENTVNIETYYNLMRSTASGLAFLRNQIGGSK
subdomains  V                VI-A

cons.aa      DLK N          DFG
htGFBR-II   -GRPKMPIVHRDLKSSNIVKNDLTCCLDFGLSLRL---GPYSSVDDLANSQGVGTARYMAP
mActr-IIB   GEGHKPSIAHRDFKSKNVLLKSDLTAVLADFLGLAVRF---EPGKPPGD--THGQVGTTRYMAP
mActr-II    -DGHKPAISHRDIKSKNVLLKNNLTACIADFLGLALRF---EAGKSAGD--THGQVGTTRYMAP
daf-1       -ESNKPAMAHRDIKSKNIMYKNDLTCAIGDLGLSLSKPEDAASDI IAN--ENYKCGTVRYLAP
subdomains  VI-B          VII          VIII
    
```

(57) Abstract

A new receptor family has been identified, of activin-like kinases. Novel proteins have activin/TGF-β-type I receptor functionality, and have consequential diagnostic/therapeutic utility. They may have a serine/threonine kinase domain, a DFKSRN or DLKSKN sequence in subdomain VIB and/or a GTKRYM sequence in subdomain VIII.

BEST AVAILABLE COPY

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LJ	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TC	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

ACTIVIN RECEPTOR-LIKE KINASES, PROTEINS HAVING
SERINE THREONINE KINASE DOMAINS AND THEIR USE.

Field of the Invention

5 This invention relates to proteins having serine/threonine kinase domains, corresponding nucleic acid molecules, and their use.

Background of the Invention

10 The transforming growth factor- β (TGF- β) superfamily consists of a family of structurally-related proteins, including three different mammalian isoforms of TGF- β (TGF- β 1, β 2 and β 3), activins, inhibins, müllerian-inhibiting substance and bone morphogenic proteins (BMPs) (for reviews see Roberts and Sporn, (1990) Peptide Growth Factors and Their Receptors, Pt.1, Sporn and Roberts, eds. (Berlin: Springer - Verlag) pp 419-472; Moses et al (1990) Cell 63, 245-247). The proteins of the TGF- β superfamily have a wide variety of biological activities. TGF- β acts as a growth inhibitor for many cell types and appears to play a central role in the regulation of embryonic development, tissue regeneration, immuno-regulation, as well as in fibrosis and carcinogenesis (Roberts and Sporn (199) see above).

25 Activins and inhibins were originally identified as factors which regulate secretion of follicle-stimulating hormone secretion (Vale et al (1990) Peptide Growth Factors and Their Receptors, Pt.2, Sporn and Roberts, eds. (Berlin: Springer-Verlag) pp.211-248). Activins were also shown to induce the differentiation of haematopoietic progenitor cells (Murata et al (1988) Proc. Natl. Acad. Sci. USA 85, 2434 - 2438; Eto et al (1987) Biochem. Biophys. Res. Commun. 142, 1095-1103) and induce mesoderm formation in Xenopus embryos (Smith et al (1990) Nature 345, 729-731; van den Eijnden-Van Raaij et al (1990) Nature 345, 732-734).

35 BMPs or osteogenic proteins which induce the formation of bone and cartilage when implanted subcutaneously (Wozney et al (1988) Science 242, 1528-1534), facilitate neuronal

differentiation (Paralkar et al (1992) J. Cell Biol. 119, 1721-1728) and induce monocyte chemotaxis (Cunningham et al (1992) Proc. Natl. Acad. Sci. USA 89, 11740-11744). Müllerian-inhibiting substance induces regression of the Müllerian duct in the male reproductive system (Cate et al (1986) Cell 45, 685-698), and a glial cell line-derived neurotrophic factor enhances survival of midbrain dopaminergic neurons (Lin et al (1993) Science 260, 1130-1132). The action of these growth factors is mediated through binding to specific cell surface receptors.

Within this family, TGF- β receptors have been most thoroughly characterized. By covalently cross-linking radio-labelled TGF- β to cell surface molecules followed by polyacrylamide gel electrophoresis of the affinity-labelled complexes, three distinct size classes of cell surface proteins (in most cases) have been identified, denoted receptor type I (53 kd), type II (75 kd), type III or betaglycan (a 300 kd proteoglycan with a 120 kd core protein) (for a review see Massague (1992) Cell 69 1067-1070) and more recently endoglin (a homodimer of two 95 kd subunits) (Cheifetz et al (1992) J. Biol. Chem. 267 19027-19030). Current evidence suggests that type I and type II receptors are directly involved in receptor signal transduction (Segarini et al (1989) Mol. Endo., 3, 261-272; Laiho et al (1991) J. Biol. Chem. 266, 9100-9112) and may form a heteromeric complex; the type II receptor is needed for the binding of TGF- β to the type I receptor and the type I receptor is needed for the signal transduction induced by the type II receptor (Wrana et al (1992) Cell, 71, 1003-1004). The type III receptor and endoglin may have more indirect roles, possibly by facilitating the binding of ligand to type II receptors (Wang et al (1991) Cell, 67 797-805; López-Casillas et al (1993) Cell, 73 1435-1444).

Binding analyses with activin A and BMP4 have led to the identification of two co-existing cross-linked affinity complexes of 50-60 kDa and 70-80 kDa on responsive cells

(Hino et al (1989) J. Biol. Chem. 264, 10309 - 10314; Mathews and Vale (1991), Cell 68, 775-785; Paralker et al (1991) Proc. Natl. Acad. Sci. USA 87, 8913-8917). By analogy with TGF- β receptors they are thought to be signalling receptors and have been named type I and type II receptors.

Among the type II receptors for the TGF- β superfamily of proteins, the cDNA for the activin type II receptor (Act RII) was the first to be cloned (Mathews and Vale (1991) Cell 65, 973-982). The predicted structure of the receptor was shown to be a transmembrane protein with an intracellular serine/threonine kinase domain. The activin receptor is related to the C. elegans daf-1 gene product, but the ligand is currently unknown (Georgi et al (1990) Cell 61, 635-645). Thereafter, another form of the activin type II receptor (activin type IIB receptor), of which there are different splicing variants (Mathews et al (1992), Science 225, 1702-1705; Attisano et al (1992) Cell 68, 97-108), and the TGF- β type II receptor (TBR II) (Lin et al (1992) Cell 68, 775-785) were cloned, both of which have putative serine/threonine kinase domains.

Summary of the Invention

The present invention involves the discovery of related novel peptides, including peptides having the activity of those defined herein as SEQ ID Nos. 2, 4, 8, 10, 12, 14, 16 and 18. Their discovery is based on the realisation that receptor serine/threonine kinases form a new receptor family, which may include the type II receptors for other proteins in the TGF- β superfamily. To ascertain whether there were other members of this family of receptors, a protocol was designed to clone ActRII/daf I related cDNAs. This approach made use of the polymerase chain reaction (PCR), using degenerate primers based upon the amino-acid sequence similarity between kinase domains of the mouse activin type II receptor and daf-I gene products.

This strategy resulted in the isolation of a new family of receptor kinases called Activin receptor like kinases (ALK's) 1-6. These cDNAs showed an overall 33-39% sequence similarity with ActRII and TGF- β type II receptor and 40-92% sequence similarity towards each other in the kinase domains.

Soluble receptors according to the invention comprise at least predominantly the extracellular domain. These can be selected from the information provided herein, prepared in conventional manner, and used in any manner associated with the invention.

Antibodies to the peptides described herein may be raised in conventional manner. By selecting unique sequences of the peptides, antibodies having desired specificity can be obtained.

The antibodies may be monoclonal, prepared in known manner. In particular, monoclonal antibodies to the extracellular domain are of potential value in therapy.

Products of the invention are useful in diagnostic methods, e.g. to determine the presence in a sample for an analyte binding therewith, such as in an antagonist assay. Conventional techniques, e.g. an enzyme-linked immunosorbent assay, may be used.

Products of the invention having a specific receptor activity can be used in therapy, e.g. to modulate conditions associated with activin or TGF- β activity. Such conditions include fibrosis, e.g. liver cirrhosis and pulmonary fibrosis, cancer, rheumatoid arthritis and glomeronephritis.

Brief Description of the Drawings

Figure 1 shows the alignment of the serine/threonine (S/T) kinase domains (I-VIII) of related receptors from transmembrane proteins, including embodiments of the present invention. The nomenclature of the subdomains is accordingly to Hanks et al (1988).

Figures 2A to 2D shows the sequences and characteristics of the respective primers used in the

initial PCR reactions. The nucleic acid sequences are also given as SEQ ID Nos. 19 to 22.

Figure 3 is a comparison of the amino-acid sequences of human activin type II receptor (Act R-II), mouse activin type IIB receptor (Act R-IIB), human TGF- β type II receptor (TBR-II), human TGF- β type I receptor (ALK-5), human activin receptor type IA (ALK-2), and type IB (ALK-4), ALKs 1 & 3 and mouse ALK-6.

Figure 4 shows, schematically, the structures for Daf-1, Act R-II, Act R-IIB, TBR-II, TBR-I/ALK-5, ALK's -1, -2 (Act RIA), -3, -4 (Act RIB) & -6.

Figure 5 shows the sequence alignment of the cysteine-rich domains of the ALKs, TBR-II, Act R-II, Act R-IIB and daf-1 receptors.

Figure 6 is a comparison of kinase domains of serine/threonine kinases, showing the percentage amino-acid identity of the kinase domains.

Figure 7 shows the pairwise alignment relationship between the kinase domains of the receptor serine/threonine kinases. The dendrogram was generated using the Jotun-Hein alignment program (Hein (1990) Meth. Enzymol. 183, 626-645).

Brief Description of the Sequence Listings

Sequences 1 and 2 are the nucleotide and deduced amino-acid sequences of cDNA for hALK-1 (clone HP57).

Sequences 3 and 4 are the nucleotide and deduced amino-acid sequences of cDNA for hALK-2 (clone HP53).

Sequences 5 and 6 are the nucleotide and deduced amino-acid sequences of cDNA for hALK-3 (clone ONF5).

Sequences 7 and 8 the nucleotide and deduced amino-acid sequences of cDNA for hALK-4 (clone 11H8), complemented with PCR product encoding extracellular domain.

Sequences 9 and 10 are the nucleotide and deduced amino-acid sequences of cDNA for hALK-5 (clone EMBLA).

Sequences 11 and 12 are the nucleotide and deduced amino-acid sequences of cDNA for mALK-1 (clone AM6).

Sequences 13 and 14 are the nucleotide and deduced amino-acid sequences of cDNA for mALK-3 (clones ME-7 and ME-D).

5 Sequences 15 and 16 are the nucleotide and deduced amino-acid sequences of cDNA for mALK-4 (clone 8a1).

Sequences 17 and 18 are the nucleotide and deduced amino-acid sequences of cDNA for mALK-6 (clone ME-6).

10 Sequence 19 (B1-S) is a sense primer, extracellular domain, cysteine-rich region, BamHI site at 5' end, 28-mer, 64-fold degeneracy.

Sequence 20 (B3-S) is a sense primer, kinase domain II, BamHI site at 5' end, 25-mer, 162-fold degeneracy.

15 Sequence 21 (B7-S) is a sense primer, kinase domain VIB, S/T kinase specific residues, BamHI site at 5' end, 24-mer, 288-fold degeneracy.

Sequence 22 (E8-AS) is an anti-sense primer, kinase domain, S/T kinase-specific residues EcoRI site at 5' end, 20-mer, 18-fold degeneracy.

Sequence 23 is an oligonucleotide probe.

20 Sequence 24 is a 5' primer.

Sequence 25 is a 3' primer.

Sequence 26 is a consensus sequence in Subdomain I.

Sequences 27 and 28 are novel sequence motifs in Subdomain VIB.

25 Sequence 29 is a novel sequence motif in Subdomain VIII.

Description of the Invention

As described in more detail below, nucleic acid sequences have been isolated, coding for a new sub-family
30 of serine/threonine receptor kinases. The term nucleic acid molecules as used herein refers to any sequence which codes for the murine, human or mammalian form, amino-acid sequences of which are presented herein. It is understood that the well known phenomenon of codon degeneracy provides
35 for a great deal of sequence variation and all such varieties are included within the scope of this invention.

The nucleic acid sequences described herein may be used to clone the respective genomic DNA sequences in order to study the genes' structure and regulation. The murine and human cDNA or genomic sequences can also be used to
5 isolate the homologous genes from other mammalian species. The mammalian DNA sequences can be used to study the receptors' functions in various in vitro and in vivo model systems.

As exemplified below for ALK-5 cDNA, it is also
10 recognised that, given the sequence information provided herein, the artisan could easily combine the molecules with a pertinent promoter in a vector, so as to produce a cloning vehicle for expression of the molecule. The promoter and coding molecule must be operably linked via
15 any of the well-recognized and easily-practised methodologies for so doing. The resulting vectors, as well as the isolated nucleic acid molecules themselves, may be used to transform prokaryotic cells (e.g. E. coli), or transfect eukaryotes such as yeast (S. cerevisiae), PAE,
20 COS or CHO cell lines. Other appropriate expression systems will also be apparent to the skilled artisan.

Several methods may be used to isolate the ligands for the ALKs. As shown for ALK-5 cDNA, cDNA clones encoding the active open reading frames can be subcloned into
25 expression vectors and transfected into eukaryotic cells, for example COS cells. The transfected cells which can express the receptor can be subjected to binding assays for radioactively-labelled members of the TGF- β superfamily (TGF- β , activins, inhibins, bone morphogenic proteins and müllerian-inhibiting substances), as it may be expected
30 that the receptors will bind members of the TGF- β superfamily. Various biochemical or cell-based assays can be designed to identify the ligands, in tissue extracts or conditioned media, for receptors in which a ligand is not
35 known. Antibodies raised to the receptors may also be used to identify the ligands, using the immunoprecipitation of the cross-linked complexes. Alternatively, purified

receptor could be used to isolate the ligands using an affinity-based approach. The determination of the expression patterns of the receptors may also aid in the isolation of the ligand. These studies may be carried out
5 using ALK DNA or RNA sequences as probes to perform in situ hybridisation studies.

The use of various model systems or structural studies should enable the rational development of specific agonists and antagonists useful in regulating receptor function. It
10 may be envisaged that these can be peptides, mutated ligands, antibodies or other molecules able to interact with the receptors.

The foregoing provides examples of the invention Applicants intend to claim which includes, inter alia,
15 isolated nucleic acid molecules coding for activin receptor-like kinases (ALKs), as defined herein. These include such sequences isolated from mammalian species such as mouse, human, rat, rabbit and monkey.

The following description relates to specific
20 embodiments. It will be understood that the specification and examples are illustrative but not limitative of the present invention and that other embodiments within the spirit and scope of the invention will suggest themselves to those skilled in the art.

25 Preparation of mRNA and Construction of a cDNA Library

For construction of a cDNA library, poly (A)⁺ RNA was isolated from a human erythroleukemia cell line (HEL 92.1.7) obtained from the American Type Culture Collection (ATCC TIB 180). These cells were chosen as they have been
30 shown to respond to both activin and TGF- β . Moreover leukaemic cells have proved to be rich sources for the cloning of novel receptor tyrosine kinases (Partanen et al (1990) Proc. Natl. Acad. Sci. USA 87, 8913-8917 and (1992) Mol. Cell. Biol. 12, 1698-1707). (Total) RNA was prepared
35 by the guanidinium isothiocyanate method (Chirgwin et al (1979) Biochemistry 18, 5294-5299). mRNA was selected using the poly-A or poly AT tract mRNA isolation kit

(Promega, Madison, Wisconsin, U.S.A.) as described by the manufacturers, or purified through an oligo (dT)-cellulose column as described by Aviv and Leder (1972) Proc. Natl. Acad. Sci. USA 69, 1408-1412. The isolated mRNA was used for the synthesis of random primed (Amersham) cDNA, that was used to make a λ gt10 library with 1×10^5 independent cDNA clones using the Riboclone cDNA synthesis system (Promega) and λ gt10 in vitro packaging kit (Amersham) according to the manufacturers' procedures. An amplified oligo (dT) primed human placenta λ ZAPII cDNA library of 5×10^5 independent clones was used. Poly (A)⁺ RNA isolated from AG1518 human foreskin fibroblasts was used to prepare a primary random primed λ ZAPII cDNA library of 1.5×10^6 independent clones using the RiboClone cDNA synthesis system and Gigapack Gold II packaging extract (Stratagene). In addition, a primary oligo (dT) primed human foreskin fibroblast λ gt10 cDNA library (Claesson-Welsh et al (1989) Proc. Natl. Acad. Sci. USA. 86 4917-4912) was prepared. An amplified oligo (dT) primed HEL cell λ gt11 cDNA library of 1.5×10^6 independent clones (Poncz et al (1987) Blood 69 219-223) was used. A twelve-day mouse embryo λ EXIox cDNA library was obtained from Novagen (Madison, Wisconsin, U.S.A.); a mouse placenta λ ZAPII cDNA library was also used.

25 Generation of cDNA Probes by PCR

For the generation of cDNA probes by PCR (Lee et al (1988) Science 239, 1288-1291) degenerate PCR primers were constructed based upon the amino-acid sequence similarity between the mouse activin type II receptor (Mathews and Vale (1991) Cell 65, 973-982) and daf-1 (George et al (1990) Cell 61, 635-645) in the kinase domains II and VIII. Figure 1 shows the aligned serine/threonine kinase domains (I-VIII), of four related receptors of the TGF- β superfamily, i.e. hT β R-II, mActR-IIB, mActR-II and the daf-1 gene product, using the nomenclature of the subdomains according to Hanks et al (1988) Science 241, 45-52.

Several considerations were applied in the design of the PCR primers. The sequences were taken from regions of homology between the activin type II receptor and the daf-1 gene product, with particular emphasis on residues that confer serine/threonine specificity (see Table 2) and on residues that are shared by transmembrane kinase proteins and not by cytoplasmic kinases. The primers were designed so that each primer of a PCR set had an approximately similar GC composition, and so that self complementarity and complementarity between the 3' ends of the primer sets were avoided. Degeneracy of the primers was kept as low as possible, in particular avoiding serine, leucine and arginine residues (6 possible codons), and human codon preference was applied. Degeneracy was particularly avoided at the 3' end as, unlike the 5' end, where mismatches are tolerated, mismatches at the 3' end dramatically reduce the efficiency of PCR.

In order to facilitate directional subcloning, restriction enzyme sites were included at the 5' end of the primers, with a GC clamp, which permits efficient restriction enzyme digestion. The primers utilised are shown in Figure 2. Oligonucleotides were synthesized using Gene assembler plus (Pharmacia - LKB) according to the manufacturers instructions.

The mRNA prepared from HEL cells as described above was reverse-transcribed into cDNA in the presence of 50 mM Tris-HCl, pH 8.3, 8 mM MgCl₂, 30 mM KCl, 10 mM dithiothreitol, 2mM nucleotide triphosphates, excess oligo (dT) primers and 34 units of AMV reverse transcriptase at 42°C for 2 hours in 40 µl of reaction volume. Amplification by PCR was carried out with a 7.5% aliquot (3 µl) of the reverse-transcribed mRNA, in the presence of 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 M MgCl₂, 0.01% gelatin, 0.2 mM nucleotide triphosphates, 1 µM of both sense and antisense primers and 2.5 units of Taq polymerase (Perkin Elmer Cetus) in 100 µl reaction volume. Amplifications were performed on a thermal cycler (Perkin Elmer Cetus)

using the following program: first 5 thermal cycles with denaturation for 1 minute at 94°C, annealing for 1 minute at 50°C, a 2 minute ramp to 55°C and elongation for 1 minute at 72°C, followed by 20 cycles of 1 minute at 94°C, 30
5 seconds at 55°C and 1 minute at 72°C. A second round of PCR was performed with 3 µl of the first reaction as a template. This involved 25 thermal cycles, each composed of 94°C (1 min), 55°C (0.5 min), 72°C (1 min).

General procedures such as purification of nucleic
10 acids, restriction enzyme digestion, gel electrophoresis, transfer of nucleic acid to solid supports and subcloning were performed essentially according to established procedures as described by Sambrook *et al*, (1989), Molecular cloning: A Laboratory Manual, 2nd Ed. Cold Spring
15 Harbor Laboratory (Cold Spring Harbor, New York, USA).

Samples of the PCR products were digested with BamHI and EcoRI and subsequently fractionated by low melting point agarose gel electrophoresis. Bands corresponding to the approximate expected sizes, (see Table 1: ≈460 bp for
20 primer pair B3-S and E8-AS and ≈ 140 bp for primer pair B7-S and E8-AS) were excised from the gel and the DNA was purified. Subsequently, these fragments were ligated into pUC19 (Yanisch-Perron *et al* (1985) Gene 33, 103-119), which had been previously linearised with BamHI and EcoRI and
25 transformed into *E. coli* strain DH5α using standard protocols (Sambrook *et al*, *supra*). Individual clones were sequenced using standard double-stranded sequencing techniques and the dideoxynucleotide chain termination method as described by Sanger *et al* (1977) Proc. Natl.
30 Acad. Sci. USA 74, 5463-5467, and T7 DNA polymerase.

Employing Reverse Transcriptase PCR on HEL mRNA with the primer pair B3-S and E8-AS, three PCR products were obtained, termed 11.1, 11.2 and 11.3, that corresponded to novel genes. Using the primer pair B7-S and E8-AS, an
35 additional novel PCR product was obtained termed 5.2.

TABLE 1

5	NAME OF PCR PRODUCT	PRIMERS	INSERT SIZE (bp)	SIZE OF DNA FRAGMENT IN λ ACTRII/hTBRII CLONES (bp)	SEQUENCE IDENTITY WITH SEQUENCE λ ACTRII/hTBRII (%)	SEQUENCE IDENTITY BETWEEN λ ACTRII and TBR-II (%)
	11.1	B3-S/E8-AS	460	460	46/40	42
	11.2	B3-S/E8-AS	460	460	49/44	47
10	11.3	B3-S/E8-AS	460	460	44/36	48
	11.29	B3-S/E8-AS	460	460	ND/100	ND
	9.2	B1-S/E8-AS	800	795	100/ND	ND
	5.2	B7-S/E8-AS	140	143	40/38	60

15 Isolation of cDNA Clones

The PCR products obtained were used to screen various cDNA libraries described supra. Labelling of the inserts of PCR products was performed using random priming method (Feinberg and Vogelstein (1983) Anal. Biochem, 132 6-13) using the Megaprime DNA labelling system (Amersham). The oligonucleotide derived from the sequence of the PCR product 5.2 was labelled by phosphorylation with T4 polynucleotide kinase following standard protocols (Sambrook et al, supra). Hybridization and purification of positive bacteriophages were performed using standard molecular biological techniques.

The double-stranded DNA clones were all sequenced using the dideoxynucleotide chain-termination method as described by Sanger et al, supra, using T7 DNA polymerase (Pharmacia - LKB) or Sequenase (U.S. Biochemical Corporation, Cleveland, Ohio, U.S.A.). Compressions of nucleotides were resolved using 7-deaza-GTP (U.S. Biochemical Corp.) DNA sequences were analyzed using the DNA STAR computer program (DNA STAR Ltd. U.K.). Analyses of the sequences obtained revealed the existence of six

distinct putative receptor serine/threonine kinases which have been named ALK 1-6.

To clone cDNA for ALK-1 the oligo (dT) primed human placenta cDNA library was screened with a radiolabelled insert derived from the PCR product 11.3; based upon their restriction enzyme digestion patterns, three different types of clones with approximate insert sizes. of 1.7 kb, 2 kb & 3.5 kb were identified. The 2 kb clone, named HP57, was chosen as representative of this class and subjected to complete sequencing. Sequence analysis of ALK-1 revealed a sequence of 1984 nucleotides including a poly-A tail (SEQ ID No. 1). The longest open reading frame encodes a protein of 503 amino-acids, with high sequence similarity to receptor serine/threonine kinases (see below). The first methionine codon, the putative translation start site, is at nucleotide 283-285 and is preceded by an in-frame stop codon. This first ATG is in a more favourable context for translation initiation (Kozak (1987) Nucl. Acids Res., 15, 8125-8148) than the second and third in-frame ATG at nucleotides 316-318 and 325-327. The putative initiation codon is preceded by a 5' untranslated sequence of 282 nucleotides that is GC-rich (80% GC), which is not uncommon for growth factor receptors (Kozak (1991) J. Cell Biol., 115, 887-903). The 3' untranslated sequence comprises 193 nucleotides and ends with a poly-A tail. No bona fide poly-A addition signal is found, but there is a sequence (AATACA), 17-22 nucleotides upstream of the poly-A tail, which may serve as a poly-A addition signal.

ALK-2 cDNA was cloned by screening an amplified oligo (dT) primed human placenta cDNA library with a radiolabelled insert derived from the PCR product 11.2. Two clones, termed HP53 and HP64, with insert sizes of 2.7 kb and 2.4 kb respectively, were identified and their sequences were determined. No sequence difference in the overlapping clones was found, suggesting they are both derived from transcripts of the same gene.

Sequence analysis of cDNA clone HP53 (SEQ ID No. 3) revealed a sequence of 2719 nucleotides with a poly-A tail. The longest open reading frame encodes a protein of 509 amino-acids. The first ATG at nucleotides 104-106 agrees favourably with Kozak's consensus sequence with an A at position 3. This ATG is preceded in-frame by a stop codon. There are four ATG codons in close proximity further downstream, which agree with the Kozak's consensus sequence (Kozak, supra), but according to Kozak's scanning model the first ATG is predicted to be the translation start site. The 5' untranslated sequence is 103 nucleotides. The 3' untranslated sequence of 1089 nucleotides contains a polyadenylation signal located 9-14 nucleotides upstream from the poly-A tail. The cDNA clone HP64 lacks 498 nucleotides from the 5' end compared to HP53, but the sequence extended at the 3' end with 190 nucleotides and poly-A tail is absent. This suggests that different polyadenylation sites occur for ALK-2. In Northern blots, however, only one transcript was detected (see below).

The cDNA for human ALK-3 was cloned by initially screening an oligo (dT) primed human foreskin fibroblast cDNA library with an oligonucleotide (SEQ ID No. 23) derived from the PCR product 5.2. One positive cDNA clone with an insert size of 3 kb, termed ON11, was identified. However, upon partial sequencing, it appeared that this clone was incomplete; it encodes only part of the kinase domain and lacks the extracellular domain. The most 5' sequence of ON11, a 540 nucleotide XbaI restriction fragment encoding a truncated kinase domain, was subsequently used to probe a random primed fibroblast cDNA library from which one cDNA clone with an insert size of 3 kb, termed ONF5, was isolated (SEQ ID No. 5). Sequence analysis of ONF5 revealed a sequence of 2932 nucleotides without a poly-A tail, suggesting that this clone was derived by internal priming. The longest open reading frame codes for a protein of 532 amino-acids. The first ATG codon which is compatible with Kozak's consensus

sequence (Kozak, supra), is at 310-312 nucleotides and is preceded by an in-frame stop codon. The 5' and 3' untranslated sequences are 309 and 1027 nucleotides long, respectively.

5 ALK-4 cDNA was identified by screening a human oligo (dT) primed human erythroleukemia cDNA library with the radiolabelled insert of the PCR product 11.1 as a probe. One cDNA clone, termed 11H8, was identified with an insert size of 2 kb (SEQ ID No. 7). An open reading frame was
10 found encoding a protein sequence of 383 amino-acids encoding a truncated extracellular domain with high similarity to receptor serine/threonine kinases. The 3' untranslated sequence is 818 nucleotides and does not contain a poly-A tail, suggesting that the cDNA was
15 internally primed. cDNA encoding the complete extracellular domain (nucleotides 1-366) was obtained from HEL cells by RT-PCR with 5' primer (SEQ ID No. 24) derived in part from sequence at translation start site of SKR-2 (a cDNA sequence deposited in GenBank data base, accession
20 number L10125, that is identical in part to ALK-4) and 3' primer (SEQ ID No. 25) derived from 11H8 cDNA clone.

ALK-5 was identified by screening the random primed HEL cell λ gt 10 cDNA library with the PCR product 11.1 as a probe. This yielded one positive clone termed EMBLA
25 (insert size of 5.3 kb with 2 internal EcoRI sites). Nucleotide sequencing revealed an open reading frame of 1509 bp, coding for 503 amino-acids. The open reading frame was flanked by a 5' untranslated sequence of 76 bp, and a 3' untranslated sequence of 3.7 kb which was not
30 completely sequenced. The nucleotide and deduced amino-acid sequences of ALK-5 are shown in SEQ ID Nos. 9 and 10. In the 5' part of the open reading frame, only one ATG codon was found; this codon fulfils the rules of translation initiation (Kozak, supra). An in-frame stop
35 codon was found at nucleotides (-54)-(-52) in the 5' untranslated region. The predicted ATG start codon is followed by a stretch of hydrophobic amino-acid residues

which has characteristics of a cleavable signal sequence. Therefore, the first ATG codon is likely to be used as a translation initiation site. A preferred cleavage site for the signal peptidase, according to von Heijne (1986) Nucl. Acid. Res. 14, 4683-4690, is located between amino-acid residues 24 and 25. The calculated molecular mass of the primary translated product of the ALK-5 without signal sequence is 53,646 Da.

Screening of the mouse embryo λ EX Iox cDNA library using PCR, product 11.1 as a probe yielded 20 positive clones. DNAs from the positive clones obtained from this library were digested with EcoRI and HindIII, electrophoretically separated on a 1.3% agarose gel and transferred to nitrocellulose filters according to established procedures as described by Sambrook et al, supra. The filters were then hybridized with specific probes for human ALK-1 (nucleotide 288-670), ALK-2 (nucleotide 1-581), ALK-3 (nucleotide 79-824) or ALK-4 (nucleotide 1178-1967). Such analyses revealed that a clone termed ME-7 hybridised with the human ALK-3 probe. However, nucleotide sequencing revealed that this clone was incomplete, and lacked the 5' part of the translated region. Screening the same cDNA library with a probe corresponding to the extracellular domain of human ALK-3 (nucleotides 79-824) revealed the clone ME-D. This clone was isolated and the sequence was analyzed. Although this clone was incomplete in the 3' end of the translated region, ME-7 and ME-D overlapped and together covered the complete sequence of mouse ALK-3. The predicted amino-acid sequence of mouse ALK-3 is very similar to the human sequence; only 8 amino-acid residues differ (98% identity; see SEQ ID No. 14) and the calculated molecular mass of the primary translated product without the putative signal sequence is 57,447 Da.

Of the clones obtained from the initial library screening with PCR product 11.1, four clones hybridized to the probe corresponding to the conserved kinase domain of

ALK-4 but not to probes from more divergent parts of ALK-1 to -4. Analysis of these clones revealed that they have an identical sequence which differs from those of ALK-1 to -5 and was termed ALK-6. The longest clone ME6 with a 2.0 kb insert was completely sequenced yielding a 1952 bp fragment consisting of an open reading frame of 1506 bp (502 amino-acids), flanked by a 5' untranslated sequence of 186 bp, and a 3' untranslated sequence of 160 bp. The nucleotide and predicted amino-acid sequences of mouse ALK-6 are shown in SEQ ID Nos. 17 and 18. No polyadenylation signal was found in the 3' untranslated region of ME6, indicating that the cDNA was internally primed in the 3' end. Only one ATG codon was found in the 5' part of the open reading frame, which fulfils the rules for translation initiation (Kozak, supra), and was preceded by an in-frame stop codon at nucleotides 163-165. However, a typical hydrophobic leader sequence was not observed at the N terminus of the translated region. Since there is no ATG codon and putative hydrophobic leader sequence, this ATG codon is likely to be used as a translation initiation site. The calculated molecular mass of the primary translated product with the putative signal sequence is 55,576 Da.

Mouse ALK-1 (clone AM6 with 1.9 kb insert) was obtained from the mouse placenta λ ZAPII cDNA library using human ALK-1 cDNA as a probe (see SEQ ID No. 11). Mouse ALK-4 (clone 8a1 with 2.3kb insert) was also obtained from this library using human ALK-4 cDNA library as a probe (SEQ ID No. 15).

To summarise, clones HP22, HP57, ONF1, ONF3, ONF4 and HP29 encode the same gene, ALK-1. Clone AM6 encodes mouse ALK-1. HP53, HP64 and HP84 encode the same gene, ALK-2. ONF5, ONF2 and ON11 encode the same gene ALK-3. ME-7 and ME-D encode the mouse counterpart of human ALK-3. 11H8 encodes a different gene ALK-4, whilst 8a1 encodes the mouse equivalent. EMBLA encodes ALK-5, and ME-6 encodes ALK-6.

The sequence alignment between the 6 ALK genes and TBR-II, mActR-II and ActR-IIB is shown in Figure 3. These molecules have a similar domain structure; an N-terminal predicted hydrophobic signal sequence (von Heijne (1986) Nucl. Acids Res. 14: 4683-4690) is followed by a relatively small extracellular cysteine-rich ligand binding domain, a single hydrophobic transmembrane region (Kyte & Doolittle (1982) J. Mol. Biol. 157, 105-132) and a C-terminal intracellular portion, which consists almost entirely of a kinase domain (Figures 3 and 4).

The extracellular domains of these receptors have cysteine-rich regions, but they show little sequence similarity; for example, less than 20% sequence identity is found between Daf-1, ActR-II, TBR-II and ALK-5. The ALKs appear to form a subfamily as they show higher sequence similarities (15-47% identity) in their extracellular domains. The extracellular domains of ALK-5 and ALK-4 have about 29% sequence identity. In addition, ALK-3 and ALK-6 share a high degree of sequence similarity in their extracellular domains (46% identity).

The positions of many of the cysteine residues in all receptors can be aligned, suggesting that the extracellular domains may adopt a similar structural configuration. See Figure 5 for ALKs-1,-2,-3 & -5. Each of the ALKs (except ALK-6) has a potential N-linked glycosylation site, the position of which is conserved between ALK-1 and ALK-2, and between ALK-3, ALK-4 and ALK-5 (see Figure 4).

The sequence similarities in the kinase domains between daf-1, ActR-II, TBR-II and ALK-5 are approximately 40%, whereas the sequence similarity between the ALKs 1 to 6 is higher (between 59% and 90%; see Figure 6). Pairwise comparison using the Jutun-Hein sequence alignment program (Hein (1990) Meth, Enzymol., 183, 626-645), between all family members, identifies the ALKs as a separate subclass among serine/threonine kinases (Figure 7).

The catalytic domains of kinases can be divided into 12 subdomains with stretches of conserved amino-acid

residues. The key motifs are found in serine/threonine kinase receptors suggesting that they are functional kinases. The consensus sequence for the binding of ATP (Gly-X-Gly-X-X-Gly in subdomain I followed by a Lys residue further downstream in subdomain II) is found in all the ALKs.

The kinase domains of daf-1, ActR-II, and ALKs show approximately equal sequence similarity with tyrosine and serine/threonine protein kinases. However analysis of the amino-acid sequences in subdomains VI and VIII, which are the most useful to distinguish a specificity for phosphorylation of tyrosine residues versus serine/threonine residues (Hanks et al (1988) Science 241 42-52) indicates that these kinases are serine/threonine kinases; refer to Table 2.

TABLE 2

KINASE	SUBDOMAINS	
	VIB	VIII
Serine/threonine kinase consensus	DLKPEN	G (T/S) XX (Y/F) X
5 Tyrosine kinase consensus	DLAARN	XP(I/V) (K/R) W (T/M)
Act R-II	DIKSKN	GTRRYM
Act R-IIB	DFKSKN	GTRRYM
TBR-II	DLKSSN	GTARYM
ALK-I	DFKSRN	GTKRYM
10 ALK -2, -3, -4, -5, & -6	DLKSKN	GTKRYM

The sequence motifs DLKSKN (Subdomain VIB) and GTKRYM (Subdomain VIII), that are found in most of the serine/threonine kinase receptors, agree well with the consensus sequences for all protein serine/threonine kinase receptors in these regions. In addition, these receptors, except for ALK-1, do not have a tyrosine residue surrounded by acidic residues between subdomains VII and VIII, which is common for tyrosine kinases. A unique characteristic of the members of the ALK serine/threonine kinase receptor family is the presence of two short inserts in the kinase

domain between subdomains VIA and VIB and between subdomains X and XI. In the intracellular domain, these regions, together with the juxtamembrane part and C-terminal tail, are the most divergent between family members (see Figures 3 and 4). Based on the sequence similarity with the type II receptors for TGF- β and activin, the C termini of the kinase domains of ALKs -1 to -6 are set at Ser-495, Ser-501, Ser-527, Gln-500, Gln-498 and Ser-497, respectively.

10 mRNA Expression

The distribution of ALK-1, -2, -3, -4 was determined by Northern blot analysis. A Northern blot filter with mRNAs from different human tissues was obtained from Clontech (Palo Alto, C.A.). The filters were hybridized with 32 P-labelled probes at 42°C overnight in 50% formaldehyde, 5 x standard saline citrate (SSC; 1xSSC is 50mM sodium citrate, pH 7.0, 150 mM NaCl), 0.1% SDS, 50 mM sodium phosphate, 5 x Denhardt's solution and 0.1 mg/ml salmon sperm DNA. In order to minimize cross-hybridization, probes were used that did not encode part of the kinase domains, but corresponded to the highly diverged sequences of either 5' untranslated and ligand-binding regions (probes for ALK-1, -2 and -3) or 3' untranslated sequences (probe for ALK-4). The probes were labelled by random priming using the Multiprime (or Mega-prime) DNA labelling system and [α - 32 P] dCTP (Feinberg & Vogelstein (1983) Anal. Biochem. 132: 6-13). Unincorporated label was removed by Sephadex G-25 chromatography. Filters were washed at 65°C, twice for 30 minutes in 2.5 x SSC, 0.1% SDS and twice for 30 minutes in 0.3 x SSC, 0.1% SDS before being exposed to X-ray film. Stripping of blots was performed by incubation at 90-100°C in water for 20 minutes.

The ALK-5 mRNA size and distribution were determined by Northern blot analysis as above. An EcoRI fragment of 980bp of the full length ALK-5 cDNA clone, corresponding to the C-terminal part of the kinase domain and 3'

untranslated region (nucleotides 1259-2232 in SEQ ID No. 9) was used as a probe. The filter was washed twice in 0.5 x SSC, 0.1% SDS at 55°C for 15 minutes.

Using the probe for ALK-1, two transcripts of 2.2 and 4.9kb were detected. The ALK-1 expression level varied strongly between different tissues, high in placenta and lung, moderate in heart, muscle and kidney, and low (to not detectable) in brain, liver and pancreas. The relative ratios between the two transcripts were similar in most tissues; in kidney, however, there was relatively more of the 4.9 kb transcript. By reprobating the blot with a probe for ALK-2, one transcript of 4.0 kb was detected with a ubiquitous expression pattern. Expression was detected in every tissue investigated and was highest in placenta and skeletal muscle. Subsequently the blot was reprobated for ALK-3. One major transcript of 4.4 kb and a minor transcript of 7.9 kb were detected. Expression was high in skeletal muscle, in which also an additional minor transcript of 10 kb was observed. Moderate levels of ALK-3 mRNA were detected in heart, placenta, kidney and pancreas, and low (to not detectable) expression was found in brain, lung and liver. The relative ratios between the different transcripts were similar in the tested tissues, the 4.4 kb transcript being the predominant one, with the exception for brain where both transcripts were expressed at a similar level. Probing the blot with ALK-4 indicated the presence of a transcript with the estimated size of 5.2 kb and revealed an ubiquitous expression pattern. The results of Northern blot analysis using the probe for ALK-5 showed that a 5.5 kb transcript is expressed in all human tissues tested, being most abundant in placenta and least abundant in brain and heart.

The distribution of mRNA for mouse ALK-3 and -6 in various mouse tissues was also determined by Northern blot analysis. A multiple mouse tissue blot was obtained from Clontech, Palo Alto, California, U.S.A. The filter was hybridized as described above with probes for mouse ALK-3

and ALK-6. The EcoRI-PstI restriction fragment, corresponding to nucleotides 79-1100 of ALK-3, and the SacI-HpaI fragment, corresponding to nucleotides 57-720 of ALK-6, were used as probes. The filter was washed at 65°C
5 twice for 30 minutes in 2.5 x SSC, 0.1% SDS and twice for 30 minutes with 0.3 x SSC, 0.1% SDS and then subjected to autoradiography.

Using the probe for mouse ALK-3, a 1.1 kb transcript was found only in spleen. By reprobing the blot with the
10 ALK-6 specific probe, a transcript of 7.2 kb was found in brain and a weak signal was also seen in lung. No other signal was seen in the other tissues tested, i.e. heart, liver, skeletal muscle, kidney and testis.

All detected transcript sizes were different, and thus
15 no cross-reaction between mRNAs for the different ALKs was observed when the specific probes were used. This suggests that the multiple transcripts of ALK-1 and ALK-3 are coded from the same gene. The mechanism for generation of the different transcripts is unknown at present; they may be
20 formed by alternative mRNA splicing, differential polyadenylation, use of different promoters, or by a combination of these events. Differences in mRNA splicing in the regions coding for the extracellular domains may lead to the synthesis of receptors with different
25 affinities for ligands, as was shown for mActR-IIB (Attisano et al (1992) Cell 68, 97-108) or to the production of soluble binding protein.

The above experiments describe the isolation of nucleic acid sequences coding for new family of human
30 receptor kinases. The cDNA for ALK-5 was then used to determine the encoded protein size and binding properties.

Properties of the ALKs cDNA Encoded Proteins

To study the properties of the proteins encoded by the different ALK cDNAs, the cDNA for each ALK was subcloned
35 into a eukaryotic expression vector and transfected into various cell types and then subjected to immunoprecipitation using a rabbit antiserum raised against

a synthetic peptide corresponding to part of the intracellular juxtamembrane region. This region is divergent in sequence between the various serine/threonine kinase receptors. The following amino-acid residues were

5 used:

ALK-1	145-166
ALK-2	151-172
ALK-3	181-202
ALK-4	153-171
10 ALK-5	158-179
ALK-6	151-168

The rabbit antiserum against ALK-5 was designated VPN.

The peptides were synthesized with an Applied Biosystems 430A Peptide Synthesizer using t-butoxycarbonyl chemistry and purified by reversed-phase high performance liquid chromatography. The peptides were coupled to keyhole limpet haemocyanin (Calbiochem-Behring) using glutaraldehyde, as described by Guillick et al (1985) EMBO J. 4, 2869-2877. The coupled peptides were mixed with Freunds adjuvant and used to immunize rabbits.

Transient transfection of the ALK-5 cDNA

COS-1 cells (American Type Culture Collection) and the R mutant of Mv1Lu cells (for references, see below) were cultured in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum (FBS) and 100 units/ml penicillin and 50 µg/ml streptomycin in 5% CO₂ atmosphere at 37°C. The ALK-5 cDNA (nucleotides (-76) - 2232), which includes the complete coding region, was cloned in the pSV7d vector (Truett et al, (1985) DNA 4, 333-349), and used for transfection. Transfection into COS-1 cells was performed by the calcium phosphate precipitation method (Wigler et al (1979) Cell 16, 777-785). Briefly, cells were seeded into 6-well cell culture plates at a density of 5x10⁵ cells/well, and transfected the following day with 10 µg of recombinant plasmid. After overnight incubation, cells were washed three times with a buffer containing 25 mM Tris-HCl, pH 7.4, 138 mM NaCl, 5 mM KCl, 0.7 mM CaCl₂, 0.5

mM MgCl₂ and 0.6 mM Na₂HPO₄, and then incubated with Dulbecco's modified Eagle's medium containing FBS and antibiotics. Two days after transfection, the cells were metabolically labelled by incubating the cells for 6 hours
5 in methionine and cysteine-free MCDB 104 medium with 150 µCi/ml of [³⁵S]-methionine and [³⁵S]-cysteine (*in vivo* labelling mix; Amersham). After labelling, the cells were washed with 150 mM NaCl, 25 mM Tris-HCl, pH 7.4, and then solubilized with a buffer containing 20mM Tris-HCl, pH 7.4,
10 150 mM NaCl, 10 mM EDTA, 1% Triton X-100, 1% deoxycholate, 1.5% Trasylol (Bayer) and 1 mM phenylmethylsulfonylfluoride (PMSF; Sigma). After 15 minutes on ice, the cell lysates were pelleted by centrifugation, and the supernatants were then incubated with 7 µl of preimmune serum for 1.5 hours
15 at 4°C. Samples were then given 50 µl of protein A-Sepharose (Pharmacia-LKB) slurry (50% packed beads in 150 mM NaCl, 20 mM Tris-HCl, pH 7.4, 0.2% Triton X100) and incubated for 45 minutes at 4°C. The beads were spun down by centrifugation, and the supernatants (1 ml) were then
20 incubated with either 7 µl of preimmune serum or the VPN antiserum for 1.5 hours at 4°C. For blocking, 10 µg of peptide was added together with the antiserum. Immune complexes were then given 50 µl of protein A-Sepharose (Pharmacia - LKB) slurry (50% packed beads in 150 mM NaCl,
25 20mM Tris-HCl, pH 7.4, 0.2% Triton X-100) and incubated for 45 minutes at 4°C. The beads were spun down and washed four times with a washing buffer (20 mM Tris-HCl, pH 7.4, 500 mM NaCl, 1% Triton X-100, 1% deoxycholate and 0.2% SDS), followed by one wash in distilled water. The immune
30 complexes were eluted by boiling for 5 minutes in the SDS-sample buffer (100 mM Tris-HCl, pH 8.8, 0.01% bromophenol blue, 36% glycerol, 4% SDS) in the presence of 10 mM DTT, and analyzed by SDS-gel electrophoresis using 7-15% polyacrylamide gels (Blobel and Dobberstein, (1975) J.Cell
35 Biol. 67, 835-851). Gels were fixed, incubated with Amplify (Amersham) for 20 minutes, and subjected to fluorography. A component of 53Da was seen. This

component was not seen when preimmune serum was used, or when 10 μ g blocking peptide was added together with the antiserum. Moreover, it was not detectable in samples derived from untransfected COS-1 cells using either
5 preimmune serum or the antiserum.

Digestion with Endoglycosidase F

Samples immunoprecipitated with the VPN antisera obtained as described above were incubated with 0.5 U of endoglycosidase F (Boehringer Mannheim Biochemica) in a
10 buffer containing 100 mM sodium phosphate, pH 6.1, 50 mM EDTA, 1% Triton X-100, 0.1% SDS and 1% β -mercaptoethanol at 37°C for 24 hours. Samples were eluted by boiling for 5 minutes in the SDS-sample buffer, and analyzed by SDS-polyacrylamide gel electrophoresis as described above.
15 Hydrolysis of N-linked carbohydrates by endoglycosidase F shifted the 53 kDa band to 51 kDa. The extracellular domain of ALK-5 contains one potential acceptor site for N-glycosylation and the size of the deglycosylated protein is close to the predicted size of the core protein.

20 Establishment of PAE Cell Lines Expressing ALK-5

In order to investigate whether the ALK-5 cDNA encodes a receptor for TGF- β , porcine aortic endothelial (PAE) cells were transfected with an expression vector containing the ALK-5 cDNA, and analyzed for the binding of 125 I-TGF- β 1.
25 PAE cells were cultured in Ham's F-12 medium supplemented with 10% FBS and antibiotics (Miyazono et al., (1988) J. Biol. Chem. 263, 6407-6415). The ALK-5 cDNA was cloned into the cytomegalovirus (CMV)-based expression vector pcDNA I/NEO (Invitrogen), and transfected into PAE
30 cells by electroporation. After 48 hours, selection was initiated by adding Geneticin (G418 sulphate; Gibco - BRL) to the culture medium at a final concentration of 0.5 mg/ml (Westermarck et al., (1990) Proc. Natl. Acad. Sci. USA 87, 128-132). Several clones were obtained, and after analysis
35 by immunoprecipitation using the VPN antiserum, one clone denoted PAE/TBR-1 was chosen and further analyzed.

Iodination of TGF- β 1, Binding and Affinity Crosslinking

Recombinant human TGF- β 1 was iodinated using the chloramine T method according to Frolik *et al.*, (1984) J. Biol. Chem. 259, 10995-11000. Cross-linking experiments were performed as previously described (Ichijo *et al.*, (1990) Exp. Cell Res. 187, 263-269). Briefly, cells in 6-well plates were washed with binding buffer (phosphate-buffered saline containing 0.9 mM CaCl₂, 0.49 mM MgCl₂ and 1 mg/ml bovine serum albumin (BSA)), and incubated on ice in the same buffer with ¹²⁵I-TGF- β 1 in the presence or absence of excess unlabelled TGF- β 1 for 3 hours. Cells were washed and cross-linking was done in the binding buffer without BSA together with 0.28 mM disuccinimidyl suberate (DSS; Pierce Chemical Co.) for 15 minutes on ice. The cells were harvested by the addition of 1 ml of detachment buffer (10 mM Tris-HCl, pH 7.4, 1 mM EDTA, 10% glycerol, 0.3 mM PMSF). The cells were pelleted by centrifugation, then resuspended in 50 μ l of solubilization buffer (125 mM NaCl, 10 mM Tris-HCl, pH 7.4, 1 mM EDTA, 1% Triton X-100, 0.3 mM PMSF, 1% Trasylol) and incubated for 40 minutes on ice. Cells were centrifuged again and supernatants were subjected to analysis by SDS-gel electrophoresis using 4-15% polyacrylamide gels, followed by autoradiography. ¹²⁵I-TGF- β 1 formed a 70 kDa cross-linked complex in the transfected PAE cells (PAE/TBR-I cells). The size of this complex was very similar to that of the TGF- β type I receptor complex observed at lower amounts in the untransfected cells. A concomitant increase of 94 kDa TGF- β type II receptor complex could also be observed in the PAE/TBR-I cells. Components of 150-190 kDa, which may represent crosslinked complexes between the type I and type II receptors, were also observed in the PAE/TBR-I cells.

In order to determine whether the cross-linked 70 kDa complex contained the protein encoded by the ALK-5 cDNA, the affinity cross-linking was followed by immunoprecipitation using the VPN antiserum. For this,

cells in 25 cm² flasks were used. The supernatants obtained after cross-linking were incubated with 7 μ l of preimmune serum or VPN antiserum in the presence or absence of 10 μ g of peptide for 1.5h at 4°C. Immune complexes were then added to 50 μ l of protein A-Sepharose slurry and incubated for 45 minutes at 4°C. The protein A-Sepharose beads were washed four times with the washing buffer, once with distilled water, and the samples were analyzed by SDS-gel electrophoresis using 4-15% polyacrylamide gradient gels and autoradiography. A 70 kDa cross-linked complex was precipitated by the VPN antiserum in PAE/TBR-1 cells, and a weaker band of the same size was also seen in the untransfected cells, indicating that the untransfected PAE cells contained a low amount of endogenous ALK-5. The 70 kDa complex was not observed when preimmune serum was used, or when immune serum was blocked by 10 μ g of peptide. Moreover, a coprecipitated 94 kDa component could also be observed in the PAE/TBR-I cells. The latter component is likely to represent a TGF- β type II receptor complex, since an antiserum, termed DRL, which was raised against a synthetic peptide from the C-terminal part of the TGF- β type II receptor, precipitated a 94 kDa TGF- β type II receptor complex, as well as a 70 kDa type I receptor complex from PAE/TBR-I cells.

The carbohydrate contents of ALK-5 and the TGF- β type II receptor were characterized by deglycosylation using endoglycosidase F as described above and analyzed by SDS-polyacrylamide gel electrophoresis and autoradiography. The ALK-5 cross-linked complex shifted from 70 kDa to 66 kDa, whereas that of the type II receptor shifted from 94 kDa to 82 kDa. The observed larger shift of the type II receptor band compared with that of the ALK-5 band is consistent with the deglycosylation data of the type I and type II receptors on rat liver cells reported previously (Cheifetz *et al* (1988) *J. Biol. Chem.* 263, 16984-16991), and fits well with the fact that the porcine TGF- β type II receptor has two N-glycosylation sites (Lin *et al* (1992)

Cell 68, 775-785), whereas ALK-5 has only one (see SEQ ID No. 9).

Binding of TGF- β 1 to the type I receptor is known to be abolished by transient treatment of the cells with dithiothreitol (DTT) (Cheifetz and Massague (1991) J. Biol. Chem. 266, 20767-20772; Wrana *et al* (1992) Cell 71, 1003-1014). When analyzed by affinity cross-linking, binding of 125 I-TGF- β 1 to ALK-5, but not to the type II receptor, was completely abolished by DTT treatment of PAE/TBR-1 cells. Affinity cross-linking followed by immunoprecipitation by the VPN antiserum showed that neither the ALK-5 nor the type II receptor complexes was precipitated after DTT treatment, indicating that the VPN antiserum reacts only with ALK-5. The data show that the VPN antiserum recognizes a TGF- β type I receptor, and that the type I and type II receptors form a heteromeric complex.

125 I-TGF- β 1 Binding & Affinity Crosslinking of Transfected COS Cells

Transient expression plasmids of ALKs -1 to -6 and TBR-II were generated by subcloning into the pSV7d expression vector or into the pCDNA I expression vector (Invitrogen). Transient transfection of COS-1 cells and iodination of TGF- β 1 were carried out as described above. Crosslinking and immunoprecipitation were performed as described for PAE cells above.

Transfection of cDNAs for ALKs into COS-1 cells did not show any appreciable binding of 125 I-TGF β 1, consistent with the observation that type I receptors do not bind TGF- β in the absence of type II receptors. When the TBR-II cDNA was co-transfected with cDNAs for the different ALKs, type I receptor-like complexes were seen, at different levels, in each case. COS-1 cells transfected with TBR-II and ALK cDNAs were analyzed by affinity crosslinking followed by immunoprecipitation using the DRL antisera or specific antisera against ALKs. Each one of the ALKs bound 125 I-TGF- β 1 and was coimmunoprecipitated with the TBR-II complex using the DRL antiserum. Comparison of the

efficiency of the different ALKs to form heteromeric complexes with TBR-II, revealed that ALK-5 formed such complexes more efficiently than the other ALKs. The size of the crosslinked complex was larger for ALK-3 than for
5 other ALKs, consistent with its slightly larger size.

Expression of the ALK Protein in Different Cell Types

Two different approaches were used to elucidate which ALK's are physiological type I receptors for TGF- β .

10 Firstly, several cell lines were tested for the expression of the ALK proteins by cross-linking followed by immunoprecipitation using the specific antiseras against ALKs and the TGF- β type II receptor. The mink lung epithelial cell line, Mv1Lu, is widely used to provide target cells for TGF- β action and is well characterized
15 regarding TGF- β receptors (Laiho et al (1990) J. Biol. Chem. 265, 18518-18524; Laiho et al (1991) J. Biol. Chem. 266, 9108-9112). Only the VPN antiserum efficiently precipitated both type I and type II TGF- β receptors in the wild type Mv1Lu cells. The DRL antiserum also precipitated
20 components with the same size as those precipitated by the VPN antiserum. A mutant cell line (R mutant) which lacks the TGF- β type I receptor and does not respond to TGF- β (Laiho et al, supra) was also investigated by cross-linking followed by immunoprecipitation. Consistent with the
25 results obtained by Laiho et al (1990), supra the type III and type II TGF- β receptor complexes, but not the type I receptor complex, were observed by affinity crosslinking. Crosslinking followed by immunoprecipitation using the DRL antiserum revealed only the type II receptor complex,
30 whereas neither the type I nor type II receptor complexes was seen using the VPN antiserum. When the cells were metabolically labelled and subjected to immunoprecipitation using the VPN antiserum, the 53 kDa ALK-5 protein was precipitated in both the wild-type and R mutant Mv1Lu
35 cells. These results suggest that the type I receptor expressed in the R mutant is ALK-5, which has lost the affinity for binding to TGF- β after mutation.

The type I and type II TGF- β receptor complexes could be precipitated by the VPN and DRL antisera in other cell lines, including human foreskin fibroblasts (AG1518), human lung adenocarcinoma cells (A549), and human oral squamous cell carcinoma cells (HSC-2). Affinity cross-linking studies revealed multiple TGF- β type I receptor-like complexes of 70-77 kDa in these cells. These components were less efficiently competed by excess unlabelled TGF- β 1 in HSC-2 cells. Moreover, the type II receptor complex was low or not detectable in A549 and HSC-2 cells. Cross-linking followed by immunoprecipitation revealed that the VPN antiserum precipitated only the 70 kDa complex among the 70-77 kDa components. The DRL antiserum precipitated the 94 kDa type II receptor complex as well as the 70 kDa type I receptor complex in these cells, but not the putative type I receptor complexes of slightly larger sizes. These results suggest that multiple type I TGF- β receptors may exist and that the 70 kDa complex containing ALK-5 forms a heteromeric complex with the TGF- β type II receptor cloned by Lin *et al* (1992) Cell 68, 775-785, more efficiently than the other species. In rat pheochromocytoma cells (PC12) which have been reported to have no TGF- β receptor complexes by affinity cross-linking (Massagué *et al* (1990) Ann. N.Y. Acad. Sci. 593, 59-72), neither VPN nor DRL antisera precipitated the TGF- β receptor complexes. The antisera against ALKs -1 to -4 and ALK6 did not efficiently immunoprecipitate the crosslinked receptor complexes in porcine aortic endothelial (PAE) cells or human foreskin fibroblasts.

Next, it was investigated whether ALKs could restore responsiveness to TGF- β in the R mutant of Mv1Lu cells, which lack the ligand-binding ability of the TGF- β type I receptor but have intact type II receptor. Wild-type Mv1Lu cells and mutant cells were transfected with ALK cDNA and were then assayed for the production of plasminogen activator inhibitor-1 (PAI-1) which is produced as a result of TGF- β receptor activation as described previously by

Laiho et al (1991) Mol. Cell Biol. 11, 972-978. Briefly, cells were added with or without 10 ng/ml of TGF- β 1 for 2 hours in serum-free MCDB 104 without methionine. Thereafter, cultures were labelled with [³⁵S] methionine (40 μ Ci/ml) for 2 hours. The cells were removed by washing on ice once in PBS, twice in 10 mM Tris-HCl (pH 8.0), 0.5% sodium deoxycholate, 1 mM PMSF, twice in 2 mM Tris-HCl (pH 8.0), and once in PBS. Extracellular matrix proteins were extracted by scraping cells into the SDS-sample buffer containing DTT, and analyzed by SDS-gel electrophoresis followed by fluorography using Amplify. PAI-1 can be identified as a characteristic 45kDa band (Laiho et al (1991) Mol. Cell Biol. 11, 972-978). Wild-type Mv1Lu cells responded to TGF- β and produced PAI-1, whereas the R mutant clone did not, even after stimulation by TGF- β 1. Transient transfection of the ALK-5 cDNA into the R mutant clone led to the production of PAI-1 in response to the stimulation by TGF- β 1, indicating that the ALK-5 cDNA encodes a functional TGF- β type I receptor. In contrast, the R mutant cells that were transfected with other ALKs did not produce PAI-1 upon the addition of TGF- β 1.

Using similar approaches as those described above for the identification of TGF- β -binding ALKs, the ability of ALKs to bind activin in the presence of ActRII was examined. COS-1 cells were co-transfected as described above. Recombinant human activin A was iodinated using the chloramine T method (Mathews and Vale (1991) Cell 65, 973-982). Transfected COS-1 cells were analysed for binding and crosslinking of ¹²⁵I-activin A in the presence or absence of excess unlabelled activin A. The crosslinked complexes were subjected to immunoprecipitation using DRL antisera or specific ALK antisera.

All ALKs appear to bind activin A in the presence of Act R-II. This is more clearly demonstrated by affinity cross-linking followed by immunoprecipitation. ALK-2 and ALK-4 bound ¹²⁵I-activin A and were coimmunoprecipitated.

with ActR-II. Other ALKs also bound ^{125}I -activin A but with a lower efficiency compared to ALK-2 and ALK-4.

In order to investigate whether ALKs are physiological activin type I receptors, activin responsive cells were examined for the expression of endogenous activin type I receptors. Mv1Lu cells, as well as the R mutant, express both type I and type II receptors for activin, and the R mutant cells produce PAI-1 upon the addition of activin A. Mv1Lu cells were labeled with ^{125}I -activin A, cross-linked and immunoprecipitated by the antisera against ActR-II or ALKs as described above.

The type I and type II receptor complexes in Mv1Lu cells were immunoprecipitated only by the antisera against ALK-2, ALK-4 and ActR-II. Similar results were obtained using the R mutant cells. PAE cells do not bind activin because of the lack of type II receptors for activin, and so cells were transfected with a chimeric receptor, to enable them to bind activin, as described herein. A plasmid (chim A) containing the extracellular domain and C-terminal tail of Act R-II (amino-acids -19 to 116 and 465 to 494, respectively (Mathews and Vale (1991) Cell, 65, 973-982)) and the kinase domain of TBR-II (amino-acids 160-543) (Lin *et al* (1992) Cell, 68, 775-785) was constructed and transfected into pcDNA/neo (Invitrogen). PAE cells were stably transfected with the chim A plasmid by electroporation, and cells expressing the chim A protein were established as described previously. PAE/Chim A cells were then subjected to ^{125}I -activin A labelling crosslinking and immunoprecipitation as described above.

Similar to Mv1Lu cells, activin type I receptor complexes in PAE/Chim A cells were immunoprecipitated by the ALK-2 and ALK-4 antisera. These results show that both ALK-2 and ALK-4 serve as high affinity type I receptors for activin A in these cells.

ALK-1, ALK-3 and ALK-6 bind TGF- β 1 and activin A in the presence of their respective type II receptors, but the

functional consequences of the binding of the ligands remains to be elucidated.

The invention has been described by way of example only, without restriction of its scope. The invention is
5 defined by the subject matter herein, including the claims that follow the immediately following full Sequence Listings.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: Ludwig Institute for Cancer Research
- (B) STREET: St. Mary's Hospital Medical School, Norfolk Place
- (C) CITY: Paddington, London
- (E) COUNTRY: United Kingdom
- (F) POSTAL CODE (ZIP): W2 1PG

(ii) TITLE OF INVENTION: PROTEINS HAVING SERINE/THREONINE KINASE DOMAINS, CORRESPONDING NUCLEIC ACID MOLECULES, AND THEIR USE

(iii) NUMBER OF SEQUENCES: 29

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/M\$-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1984 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 283..1791

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```

AGGAAACGGT TTATTAGGAG GGAGTGGTGG AGCTGGGCCA GGCAGGAAGA CGCTGGAATA      60
AGAAACATTT TTGCTCCAGC CCCCATCCCA GTCCCGGGAG GCTGCCGCC CAGCTGCGCC      120
GAGCGAGCCC CTCCCGGCT CCAGCCCGGT CCGGGCCGC GCCGGACCCC AGCCCGCCGT      180
CCAGCGCTGG CGGTGCAACT GCGGCCCGC GGTGGAGGGG AGGTGGCCCC GGTCCGCCGA      240

```

AGGCTAGCGC CCGCCACCC GCAGAGCGGG CCCAGAGGGA CC ATG ACC TTG GGC Met Thr Leu Gly 1	294
TCC CCC AGG AAA GGC CTT CTG ATG CTG CTG ATG GCC TTG GTG ACC CAG Ser Pro Arg Lys Gly Leu Leu Met Leu Leu Met Ala Leu Val Thr Gln 5 10 15 20	342
GGA GAC CCT GTG AAG CCG TCT CGG GGC CCG CTG GTG ACC TGC ACG TGT Gly Asp Pro Val Lys Pro Ser Arg Gly Pro Leu Val Thr Cys Thr Cys 25 30 35	390
GAG AGC CCA CAT TGC AAG GGG CCT ACC TGC CGG GGG GCC TGG TGC ACA Glu Ser Pro His Cys Lys Gly Pro Thr Cys Arg Gly Ala Trp Cys Thr 40 45 50	438
GTA GTG CTG GTG CGG GAG GAG GGG AGG CAC CCC CAG GAA CAT CGG GGC Val Val Leu Val Arg Glu Glu Gly Arg His Pro Gln Glu His Arg Gly 55 60 65	486
TGC GGG AAC TTG CAC AGG GAG CTC TGC AGG GGG CGC CCC ACC GAG TTC Cys Gly Asn Leu His Arg Glu Leu Cys Arg Gly Arg Pro Thr Glu Phe 70 75 80	534
GTC AAC CAC TAC TGC TGC GAC AGC CAC CTC TGC AAC CAC AAC GTG TCC Val Asn His Tyr Cys Cys Asp Ser His Leu Cys Asn His Asn Val Ser 85 90 95 100	582
CTG GTG CTG GAG GCC ACC CAA CCT CCT TCG GAG CAG CCG GGA ACA GAT Leu Val Leu Glu Ala Thr Gln Pro Pro Ser Glu Gln Pro Gly Thr Asp 105 110 115	630
GGC CAG CTG GCC CTG ATC CTG GGC CCC GTG CTG GCC TTG CTG GCC CTG Gly Gln Leu Ala Leu Ile Leu Gly Pro Val Leu Ala Leu Leu Ala Leu 120 125 130	678
GTG GCC CTG GGT GTC CTG GGC CTG TGG CAT GTC CGA CCG AGG CAG GAG Val Ala Leu Gly Val Leu Gly Leu Trp His Val Arg Arg Arg Gln Glu 135 140 145	726
AAG CAG CGT GGC CTG CAC AGC GAG CTG GGA GAG TCC AGT CTC ATC CTG Lys Gln Arg Gly Leu His Ser Glu Leu Gly Glu Ser Ser Leu Ile Leu 150 155 160	774
AAA GCA TCT GAG CAG GGC GAC ACG ATG TTG GGG GAC CTC CTG GAC AGT Lys Ala Ser Glu Gln Gly Asp Thr Met Leu Gly Asp Leu Leu Asp Ser 165 170 175 180	822
GAC TGC ACC ACA GGG AGT GGC TCA GGG CTC CCC TTC CTG GTG CAG AGG Asp Cys Thr Thr Gly Ser Gly Ser Gly Leu Pro Phe Leu Val Gln Arg 185 190 195	870
ACA GTG GCA CCG CAG GTT GCC TTG GTG GAG TGT GTG GGA AAA GGC CGC Thr Val Ala Arg Gln Val Ala Leu Val Glu Cys Val Gly Lys Gly Arg 200 205 210	918
TAT GGC GAA GTG TGG CGG GGC TTG TGG CAC GGT GAG AGT GTG GCC GTC Tyr Gly Glu Val Trp Arg Gly Leu Trp His Gly Glu Ser Val Ala Val 215 220 225	966

AAG ATC TTC TCC TCG AGG GAT GAA CAG TCC TGG TTC CGG GAG ACT GAG	1014
Lys Ile Phe Ser Ser Arg Asp Glu Gln Ser Trp Phe Arg Glu Thr Glu	
230 235 240	
ATC TAT AAC ACA GTA TTG CTC AGA CAC GAC AAC ATC CTA GGC TTC ATC	1062
Ile Tyr Asn Thr Val Leu Leu Arg His Asp Asn Ile Leu Gly Phe Ile	
245 250 255 260	
GCC TCA GAC ATG ACC TCC CGC AAC TCG AGC ACG CAG CTG TGG CTC ATC	1110
Ala Ser Asp Met Thr Ser Arg Asn Ser Ser Thr Gln Leu Trp Leu Ile	
265 270 275	
ACG CAC TAC CAC GAG CAC GGC TCC CTC TAC GAC TTT CTG CAG AGA CAG	1158
Thr His Tyr His Glu His Gly Ser Leu Tyr Asp Phe Leu Gln Arg Gln	
280 285 290	
ACG CTG GAG CCC CAT CTG GCT CTG AGG CTA GCT GTG TCC GCG GCA TGC	1206
Thr Leu Glu Pro His Leu Ala Leu Arg Leu Ala Val Ser Ala Ala Cys	
295 300 305	
GGC CTG GCG CAC CTG CAC GTG GAG ATC TTC GGT ACA CAG GGC AAA CCA	1254
Gly Leu Ala His Leu His Val Glu Ile Phe Gly Thr Gln Gly Lys Pro	
310 315 320	
GCC ATT GCC CAC CGC GAC TTC AAG AGC CGC AAT GTG CTG GTC AAG AGC	1302
Ala Ile Ala His Arg Asp Phe Lys Ser Arg Asn Val Leu Val Lys Ser	
325 330 335 340	
AAC CTG CAG TGT TGC ATC GCC GAC CTG GGC CTG GCT GTG ATG CAC TCA	1350
Asn Leu Gln Cys Cys Ile Ala Asp Leu Gly Leu Ala Val Met His Ser	
345 350 355	
CAG GGC AGC GAT TAC CTG GAC ATC GGC AAC AAC CCG AGA GTG GGC ACC	1398
Gln Gly Ser Asp Tyr Leu Asp Ile Gly Asn Asn Pro Arg Val Gly Thr	
360 365 370	
AAG CGG TAC ATG GCA CCC GAG GTG CTG GAC GAG CAG ATC CGC ACG GAC	1446
Lys Arg Tyr Met Ala Pro Glu Val Leu Asp Glu Gln Ile Arg Thr Asp	
375 380 385	
TGC TTT GAG TCC TAC AAG TGG ACT GAC ATC TGG GCC TTT GGC CTG GTG	1494
Cys Phe Glu Ser Tyr Lys Trp Thr Asp Ile Trp Ala Phe Gly Leu Val	
390 395 400	
CTG TGG GAG ATT GCC CGC CGG ACC ATC GTG AAT GGC ATC GTG GAG GAC	1542
Leu Trp Glu Ile Ala Arg Arg Thr Ile Val Asn Gly Ile Val Glu Asp	
405 410 415 420	
TAT AGA CCA CCC TTC TAT GAT GTG GTG CCC AAT GAC CCC AGC TTT GAG	1590
Tyr Arg Pro Pro Phe Tyr Asp Val Val Pro Asn Asp Pro Ser Phe Glu	
425 430 435	
GAC ATG AAG AAG GTG GTG TGT GTG GAT CAG CAG ACC CCC ACC ATC CCT	1638
Asp Met Lys Lys Val Val Cys Val Asp Gln Gln Thr Pro Thr Ile Pro	
440 445 450	
AAC CGG CTG GCT GCA GAC CCG GTC CTC TCA GGC CTA GCT CAG ATG ATG	1686
Asn Arg Leu Ala Ala Asp Pro Val Leu Ser Gly Leu Ala Gln Met Met	
455 460 465	

CGG GAG TGC TGG TAC CCA AAC CCC TCT GCC CGA CTC ACC GCG CTG CGG	1734
Arg Glu Cys Trp Tyr Pro Asn Pro Ser Ala Arg Leu Thr Ala Leu Arg	
470 475 480	
ATC AAG AAG ACA CTA CAA AAA ATT AGC AAC AGT CCA GAG AAG CCT AAA	1782
Ile Lys Lys Thr Leu Gln Lys Ile Ser Asn Ser Pro Glu Lys Pro Lys	
485 490 495 500	
GTG ATT CAA TAGCCCAGGA GCACCTGATT CCTTTCTGCC TGCAGGGGGC	1831
Val Ile Gln	
TGGGGGGGTG GGGGGCAGTG GATGGTGCCC TATCTGGGTA GAGGTAGTGT GAGTGTGGTG	1891
TGTGCTGGGG ATGGGCAGCT GCGCCTGCCT GCTCGGCCCC CAGCCCACCC AGCCAAAAAT	1951
ACAGCTGGGC TGAAACCTGA AAAAAAAAAA AAA	1984

(2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 503 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Thr Leu Gly Ser Pro Arg Lys Gly Leu Leu Met Leu Leu Met Ala	
1 5 10 15	
Leu Val Thr Gln Gly Asp Pro Val Lys Pro Ser Arg Gly Pro Leu Val	
20 25 30	
Thr Cys Thr Cys Glu Ser Pro His Cys Lys Gly Pro Thr Cys Arg Gly	
35 40 45	
Ala Trp Cys Thr Val Val Leu Val Arg Glu Glu Gly Arg His Pro Gln	
50 55 60	
Glu His Arg Gly Cys Gly Asn Leu His Arg Glu Leu Cys Arg Gly Arg	
65 70 75 80	
Pro Thr Glu Phe Val Asn His Tyr Cys Cys Asp Ser His Leu Cys Asn	
85 90 95	
His Asn Val Ser Leu Val Leu Glu Ala Thr Gln Pro Pro Ser Glu Gln	
100 105 110	
Pro Gly Thr Asp Gly Gln Leu Ala Leu Ile Leu Gly Pro Val Leu Ala	
115 120 125	
Leu Leu Ala Leu Val Ala Leu Gly Val Leu Gly Leu Trp His Val Arg	
130 135 140	
Arg Arg Gln Glu Lys Gln Arg Gly Leu His Ser Glu Leu Gly Glu Ser	
145 150 155 160	

Ser Leu Ile Leu Lys Ala Ser Glu Gln Gly Asp Thr Met Leu Gly Asp
 165 170 175

Leu Leu Asp Ser Asp Cys Thr Thr Gly Ser Gly Ser Gly Leu Pro Phe
 180 185 190

Leu Val Gln Arg Thr Val Ala Arg Gln Val Ala Leu Val Glu Cys Val
 195 200 205

Gly Lys Gly Arg Tyr Gly Glu Val Trp Arg Gly Leu Trp His Gly Glu
 210 215 220

Ser Val Ala Val Lys Ile Phe Ser Ser Arg Asp Glu Gln Ser Trp Phe
 225 230 235 240

Arg Glu Thr Glu Ile Tyr Asn Thr Val Leu Leu Arg His Asp Asn Ile
 245 250 255

Leu Gly Phe Ile Ala Ser Asp Met Thr Ser Arg Asn Ser Ser Thr Gln
 260 265 270

Leu Trp Leu Ile Thr His Tyr His Glu His Gly Ser Leu Tyr Asp Phe
 275 280 285

Leu Gln Arg Gln Thr Leu Glu Pro His Leu Ala Leu Arg Leu Ala Val
 290 295 300

Ser Ala Ala Cys Gly Leu Ala His Leu His Val Glu Ile Phe Gly Thr
 305 310 315 320

Gln Gly Lys Pro Ala Ile Ala His Arg Asp Phe Lys Ser Arg Asn Val
 325 330 335

Leu Val Lys Ser Asn Leu Gln Cys Cys Ile Ala Asp Leu Gly Leu Ala
 340 345 350

Val Met His Ser Gln Gly Ser Asp Tyr Leu Asp Ile Gly Asn Asn Pro
 355 360 365

Arg Val Gly Thr Lys Arg Tyr Met Ala Pro Glu Val Leu Asp Glu Gln
 370 375 380

Ile Arg Thr Asp Cys Phe Glu Ser Tyr Lys Trp Thr Asp Ile Trp Ala
 385 390 395 400

Phe Gly Leu Val Leu Trp Glu Ile Ala Arg Arg Thr Ile Val Asn Gly
 405 410 415

Ile Val Glu Asp Tyr Arg Pro Pro Phe Tyr Asp Val Val Pro Asn Asp
 420 425 430

Pro Ser Phe Glu Asp Met Lys Lys Val Val Cys Val Asp Gln Gln Thr
 435 440 445

Pro Thr Ile Pro Asn Arg Leu Ala Ala Asp Pro Val Leu Ser Gly Leu
 450 455 460

Ala Gln Met Met Arg Glu Cys Trp Tyr Pro Asn Pro Ser Ala Arg Leu
 465 470 475 480

Thr Ala Leu Arg Ile Lys Lys Thr Leu Gln Lys Ile Ser Asn Ser Pro
 485 490 495
 Glu Lys Pro Lys Val Ile Gln
 500

(2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2724 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 104..1630

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

CTCCGAGTAC CCCAGTGACC AGAGTGAGAG AAGCTCTGAA CGAGGGCACG CCGCTTGAAG	60
GACTGTGGGC AGATGTGACC AAGAGCCTGC ATTAAGTTGT ACA ATG GTA GAT GGA	115
Met Val Asp Gly	
1	
GTG ATG ATT CTT CCT GTG CTT ATC ATG ATT GCT CTC CCC TCC CCT AGT	163
Val Met Ile Leu Pro Val Leu Ile Met Ile Ala Leu Pro Ser Pro Ser	
5 10 15 20	
ATG GAA GAT GAG AAG CCC AAG GTC AAC CCC AAA CTC TAC ATG TGT GTG	211
Met Glu Asp Glu Lys Pro Lys Val Asn Pro Lys Leu Tyr Met Cys Val	
25 30 35	
TGT GAA GGT CTC TCC TGC GGT AAT GAG GAC CAC TGT GAA GGC CAG CAG	259
Cys Glu Gly Leu Ser Cys Gly Asn Glu Asp His Cys Glu Gly Gln Gln	
40 45 50	
TGC TTT TCC TCA CTG AGC ATC AAC GAT GGC TTC CAC GTC TAC CAG AAA	307
Cys Phe Ser Ser Leu Ser Ile Asn Asp Gly Phe His Val Tyr Gln Lys	
55 60 65	
GGC TGC TTC CAG GTT TAT GAG CAG GGA AAG ATG ACC TGT AAG ACC CCG	355
Gly Cys Phe Gln Val Tyr Glu Gln Gly Lys Met Thr Cys Lys Thr Pro	
70 75 80	

CCG Pro 85	TCC Ser	CCT Pro	GGC Gly	CAA Gln	GCT Ala	GTG Val	GAG Glu	TGC Cys	TGC Cys	CAA Gln	GGG Gly	GAC Asp	TGG Trp	TGT Cys	AAC Asn	403
					90					95						
AGG Arg	AAC Asn	ATC Ile	ACG Thr	GCC Ala	CAG Gln	CTG Leu	CCC Pro	ACT Thr	AAA Lys	GGA Gly	AAA Lys	TCC Ser	TTC Phe	CCT Pro	GGA Gly	451
				105					110					115		
ACA Thr	CAG Gln	AAT Asn	TTC Phe	CAC His	TTG Leu	GAG Glu	GTT Val	GGC Gly	CTC Leu	ATT Ile	ATT Ile	CTC Leu	TCT Ser	GTA Val	GTG Val	499
			120					125					130			
TTC Phe	GCA Ala	GTA Val	TGT Cys	CTT Leu	TTA Leu	GCC Ala	TGC Cys	CTG Leu	CTG Leu	GGA Gly	GTT Val	GCT Ala	CTC Leu	CGA Arg	AAA Lys	547
			135				140					145				
TTT Phe	AAA Lys	AGG Arg	CGC Arg	AAC Asn	CAA Gln	GAA Glu	CGC Arg	CTC Leu	AAT Asn	CCC Pro	CGA Arg	GAC Asp	GTG Val	GAG Glu	TAT Tyr	595
			150			155					160					
GGC Gly	ACT Thr	ATC Ile	GAA Glu	GGG Gly	CTC Leu	ATC Ile	ACC Thr	ACC Thr	AAT Asn	GTT Val	GGA Gly	GAC Asp	AGC Ser	ACT Thr	TTA Leu	643
			165		170					175					180	
GCA Ala	GAT Asp	TTA Leu	TTG Leu	GAT Asp	CAT His	TCG Ser	TGT Cys	ACA Thr	TCA Ser	GGA Gly	AGT Ser	GGC Gly	TCT Ser	GGT Gly	CTT Leu	691
				185				190						195		
CCT Pro	TTT Phe	CTG Leu	GTA Val	CAA Gln	AGA Arg	ACA Thr	GTG Val	GCT Ala	CGC Arg	CAG Gln	ATT Ile	ACA Thr	CTG Leu	TTG Leu	GAG Glu	739
			200					205					210			
TGT Cys	GTC Val	GGG Gly	AAA Lys	GGC Gly	AGG Arg	TAT Tyr	GGT Gly	GAG Glu	GTG Val	TGG Trp	AGG Arg	GGC Gly	AGC Ser	TGG Trp	CAA Gln	787
		215					220					225				
GGG Gly	GAA Glu	AAT Asn	GTT Val	GCC Ala	GTG Val	AAG Lys	ATC Ile	TTC Phe	TCC Ser	TCC Ser	CGT Arg	GAT Asp	GAG Glu	AAG Lys	TCA Ser	835
			230			235					240					
TGG Trp	TTC Phe	AGG Arg	GAA Glu	ACG Thr	GAA Glu	TTG Leu	TAC Tyr	AAC Asn	ACT Thr	GTG Val	ATG Met	CTG Leu	AGG Arg	CAT His	GAA Glu	883
					250					255					260	
AAT Asn	ATC Ile	TTA Leu	GGT Gly	TTC Phe	ATT Ile	GCT Ala	TCA Ser	GAC Asp	ATG Met	ACA Thr	TCA Ser	AGA Arg	CAC His	TCC Ser	AGT Ser	931
				265					270					275		
ACC Thr	CAG Gln	CTG Leu	TGG Trp	TTA Leu	ATT Ile	ACA Thr	CAT His	TAT Tyr	CAT His	GAA Glu	ATG Met	GGA Gly	TCG Ser	TTG Leu	TAC Tyr	979
			280					285					290			
GAC Asp	TAT Tyr	CTT Leu	CAG Gln	CTT Leu	ACT Thr	ACT Thr	CTG Leu	GAT Asp	ACA Thr	GTT Val	AGC Ser	TGC Cys	CTT Leu	CGA Arg	ATA Ile	1027
		295					300					305				
GTG Val	CTG Leu	TCC Ser	ATA Ile	GCT Ala	AGT Ser	GGT Gly	CTT Leu	GCA Ala	CAT His	TTG Leu	CAC His	ATA Ile	GAG Glu	ATA Ile	TTT Phe	1075
		310				315					320					

GGG ACC CAA GGG AAA CCA GCC ATT GCC CAT CGA GAT TTA AAG AGC AAA Gly Thr Gln Gly Lys Pro Ala Ile Ala His Arg Asp Leu Lys Ser Lys 325 330 335 340	1123
AAT ATT CTG GTT AAG AAG AAT GGA CAG TGT TGC ATA GCA GAT TTG GGC Asn Ile Leu Val Lys Lys Asn Gly Gln Cys Cys Ile Ala Asp Leu Gly 345 350 355	1171
CTG GCA GTC ATG CAT TCC CAG AGC ACC AAT CAG CTT GAT GTG GGG AAC Leu Ala Val Met His Ser Gln Ser Thr Asn Gln Leu Asp Val Gly Asn 360 365 370	1219
AAT CCC CGT GTG GGC ACC AAG CGC TAC ATG GCC CCC GAA GTT CTA GAT Asn Pro Arg Val Gly Thr Lys Arg Tyr Met Ala Pro Glu Val Leu Asp 375 380 385	1267
GAA ACC ATC CAG GTG GAT TGT TTC GAT TCT TAT AAA AGG GTC GAT ATT Glu Thr Ile Gln Val Asp Cys Phe Asp Ser Tyr Lys Arg Val Asp Ile 390 395 400	1315
TGG GCC TTT GGA CTT GTT TTG TGG GAA GTG GCC AGG CCG ATG GTG AGC Trp Ala Phe Gly Leu Val Leu Trp Glu Val Ala Arg Arg Met Val Ser 405 410 415 420	1363
AAT GGT ATA GTG GAG GAT TAC AAG CCA CCG TTC TAC GAT GTG GTT CCC Asn Gly Ile Val Glu Asp Tyr Lys Pro Pro Phe Tyr Asp Val Val Pro 425 430 435	1411
AAT GAC CCA AGT TTT GAA GAT ATG AGG AAG GTA GTC TGT GTG GAT CAA Asn Asp Pro Ser Phe Glu Asp Met Arg Lys Val Val Cys Val Asp Gln 440 445 450	1459
CAA AGG CCA AAC ATA CCC AAC AGA TGG TTC TCA GAC CCG ACA TTA ACC Gln Arg Pro Asn Ile Pro Asn Arg Trp Phe Ser Asp Pro Thr Leu Thr 455 460 465	1507
TCT CTG GCC AAG CTA ATG AAA GAA TGC TGG TAT CAA AAT CCA TCC GCA Ser Leu Ala Lys Leu Met Lys Glu Cys Trp Tyr Gln Asn Pro Ser Ala 470 475 480	1555
AGA CTC ACA GCA CTG CGT ATC AAA AAG ACT TTG ACC AAA ATT GAT AAT Arg Leu Thr Ala Leu Arg Ile Lys Lys Thr Leu Thr Lys Ile Asp Asn 485 490 495 500	1603
TCC CTC GAC AAA TTG AAA ACT GAC TGT TGACATTTTC ATAGTGTCAA Ser Leu Asp Lys Leu Lys Thr Asp Cys 505	1650
GAAGGAAGAT TTGACGTTGT TGTCATTGTC CAGCTGGGAC CTAATGCTGG CCTGACTGGT	1710
TGTCAGAATG GAATCCATCT GTCTCCCTCC CCAAATGGCT GCTTTGACAA GGCAGACGTC	1770
GTACCCAGCC ATGTGTTGGG GAGACATCAA AACCACCCTA ACCTCGCTCG ATGACTGTGA	1830
ACTGGGCATT TCACGAACTG TTCACACTGC AGAGACTAAT GTTGGACAGA CACTGTTGCA	1890
AAGGTAGGGA CTGGAGGAAC ACAGAGAAAT CCTAAAAGAG ATCTGGGCAT TAAGTCAGTG	1950
GCTTTGCATA -GCTTTCACAA GTCTCCTAGA CACTCCCCAC GGGAAACTCA AGGAGGTGGT	2010

GAATTTTAA TCAGCAATAT TGCCTGTGCT TCTCTTCTTT ATTGCACTAG GAATTCTTTG 2070
 CATTCCCTTAC TTGCACTGTT ACTCTTAATT TTAAAGACCC AACTTGCCAA AATGTTGGCT 2130
 CGGTACTCCA CTGGTCTGTC TTTGGATAAT AGGAATTCAA TTTGGCAAAA CAAAATGTAA 2190
 TGTCAGACTT TGCTGCATTT TACACATGTG CTGATGTTTA CAATGATGCC GAACATTAGG 2250
 AATGTTTAT ACACAACCTT GCAAATTATT TATTACTTGT GCACTTAGTA GTTTTACAA 2310
 AACTGCTTTG TGCATATGTT AAAGCTTATT TTTATCTGGT CTTATGATTT TATTACAGAA 2370
 ATGTTTTTAA CACTATACTC TAAAATGGAC ATTTTCTTTT ATTATCAGTT AAAATCACAT 2430
 TTTAAGTGCT TCACATTGT ATGTGTGTAG ACTGTAACCT TTTTTCAGTT CATATGCAGA 2490
 ACGTATTTAG CCATTACCCA CGTGACACCA CCGAATATAT TATCGATTTA GAAGCAAAGA 2550
 TTTCAGTAGA ATTTTAGTCC TGAACGCTAC GGGGAAAATG CATTTCCTTC AGAATTATCC 2610
 ATTACGTGCA TTTAACTCT GCCAGAAAAA AATAACTATT TTGTTTTAAT CTACTTTTTC 2670
 TATTTAGTAG TTATTTGTAT AAATTAATA AACTGTTTTC AAGTCAAPAA AAAA 2724

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 509 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Val Asp Gly Val Met Ile Leu Pro Val Leu Ile Met Ile Ala Leu
 1 5 10 15
 Pro Ser Pro Ser Met Glu Asp Glu Lys Pro Lys Val Asn Pro Lys Leu
 20 25 30
 Tyr Met Cys Val Cys Glu Gly Leu Ser Cys Gly Asn Glu Asp His Cys
 35 40 45
 Glu Gly Gln Gln Cys Phe Ser Ser Leu Ser Ile Asn Asp Gly Phe His
 50 55 60
 Val Tyr Gln Lys Gly Cys Phe Gln Val Tyr Glu Gln Gly Lys Met Thr
 65 70 75 80
 Cys Lys Thr Pro Pro Ser Pro Gly Gln Ala Val Glu Cys Cys Gln Gly
 85 90 95
 Asp Trp Cys Asn Arg Asn Ile Thr Ala Gln Leu Pro Thr Lys Gly Lys
 100 105 110
 Ser Phe Pro Gly Thr Gln Asn Phe His Leu Glu Val Gly Leu Ile Ile
 115 120 125

Leu Ser Val Val Phe Ala Val Cys Leu Leu Ala Cys Leu Leu Gly Val
 130 135 140
 Ala Leu Arg Lys Phe Lys Arg Arg Asn Gln Glu Arg Leu Asn Pro Arg
 145 150 155 160
 Asp Val Glu Tyr Gly Thr Ile Glu Gly Leu Ile Thr Thr Asn Val Gly
 165 170 175
 Asp Ser Thr Leu Ala Asp Leu Leu Asp His Ser Cys Thr Ser Gly Ser
 180 185 190
 Gly Ser Gly Leu Pro Phe Leu Val Gln Arg Thr Val Ala Arg Gln Ile
 195 200 205
 Thr Leu Leu Glu Cys Val Gly Lys Gly Arg Tyr Gly Glu Val Trp Arg
 210 215 220
 Gly Ser Trp Gln Gly Glu Asn Val Ala Val Lys Ile Phe Ser Ser Arg
 225 230 235 240
 Asp Glu Lys Ser Trp Phe Arg Glu Thr Glu Leu Tyr Asn Thr Val Met
 245 250 255
 Leu Arg His Glu Asn Ile Leu Gly Phe Ile Ala Ser Asp Met Thr Ser
 260 265 270
 Arg His Ser Ser Thr Gln Leu Trp Leu Ile Thr His Tyr His Glu Met
 275 280 285
 Gly Ser Leu Tyr Asp Tyr Leu Gln Leu Thr Thr Leu Asp Thr Val Ser
 290 295 300
 Cys Leu Arg Ile Val Leu Ser Ile Ala Ser Gly Leu Ala His Leu His
 305 310 315 320
 Ile Glu Ile Phe Gly Thr Gln Gly Lys Pro Ala Ile Ala His Arg Asp
 325 330 335
 Leu Lys Ser Lys Asn Ile Leu Val Lys Lys Asn Gly Gln Cys Cys Ile
 340 345 350
 Ala Asp Leu Gly Leu Ala Val Met His Ser Gln Ser Thr Asn Gln Leu
 355 360 365
 Asp Val Gly Asn Asn Pro Arg Val Gly Thr Lys Arg Tyr Met Ala Pro
 370 375 380
 Glu Val Leu Asp Glu Thr Ile Gln Val Asp Cys Phe Asp Ser Tyr Lys
 385 390 395 400
 Arg Val Asp Ile Trp Ala Phe Gly Leu Val Leu Trp Glu Val Ala Arg
 405 410 415
 Arg Met Val Ser Asn Gly Ile Val Glu Asp Tyr Lys Pro Pro Phe Tyr
 420 425 430
 Asp Val Val Pro Asn Asp Pro Ser Phe Glu Asp Met Arg Lys Val Val
 435 440 445

Cys Val Asp Gln Gln Arg Pro Asn Ile Pro Asn Arg Trp Phe Ser Asp
 450 455 460
 Pro Thr Leu Thr Ser Leu Ala Lys Leu Met Lys Glu Cys Trp Tyr Gln
 465 470 475 480
 Asn Pro Ser Ala Arg Leu Thr Ala Leu Arg Ile Lys Lys Thr Leu Thr
 485 490 495
 Lys Ile Asp Asn Ser Leu Asp Lys Leu Lys Thr Asp Cys
 500 505

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2932 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 310..1905

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GCTCCGCGCC GAGGGCTGGA GGATGCGTTC CCTGGGGTCC GGACTTATGA AAATATGCAT	60
CAGTTTAATA CTGTCTTGGG ATTCATGAGA TGGAAGCATA GGTCAAAGCT GTTGGAGAA	120
AATCAGAAGT ACAGTTTTAT CTAGCCACAT CTTGGAGGAG TCGTAAGAAA GCAGTGGGAG	180
TTGAAGTCAT TGTCAGTGC TTGCGATCTT TTACAAGAAA ATCTCACTGA ATGATAGTCA	240
TTTAAATTGG TGAAGTAGCA AGACCAATTA TTAAGGTGA CAGTACACAG GAAACATTAC	300
AATTGAACA ATG ACT CAG CTA TAC ATT TAC ATC AGA TTA TTG GGA GCC	348
Met Thr Gln Leu Tyr Ile Tyr Ile Arg Leu Leu Gly Ala	
1 5 10	
TAT TTG TTC ATC ATT TCT CGT GTT CAA GGA CAG AAT CTG GAT AGT ATG	396
Tyr Leu Phe Ile Ile Ser Arg Val Gln Gly Gln Asn Leu Asp Ser Met	
15 20 25	

CTT Leu 30	CAT His	GGC Gly	ACT Thr	GGG Gly	ATG Met	AAA Lys	TCA Ser	CAC Asp	TCC Ser	GAC Asp	CAG Gln	AAA Lys	AAG Lys	TCA Ser	GAA Glu	444
					35					40					45	
AAT Asn	GGA Gly	GTA Val	ACC Thr	TTA Leu	GCA Ala	CCA Pro	GAG Glu	GAT Asp	ACC Thr	TTG Leu	CCT Pro	TTT Phe	TTA Leu	AAG Lys	TGC Cys	492
				50					55					60		
TAT Tyr	TGC Cys	TCA Ser	GGG Gly	CAC His	TGT Cys	CCA Pro	GAT Asp	GAT Asp	GCT Ala	ATT Ile	AAT Asn	AAC Asn	ACA Thr	TGC Cys	ATA Ile	540
			65					70					75			
ACT Thr	AAT Asn	GGA Gly	CAT His	TGC Cys	TTT Phe	GCC Ala	ATC Ile	ATA Ile	GAA Glu	GAA Glu	GAT Asp	GAC Asp	CAG Gln	GGA Gly	GAA Glu	588
		80					85					90				
ACC Thr	ACA Thr	TTA Leu	GCT Ala	TCA Ser	GGG Gly	TGT Cys	ATG Met	AAA Lys	TAT Tyr	GAA Glu	GGA Gly	TCT Ser	GAT Asp	TTT Phe	CAG Gln	636
		95				100					105					
TGC Cys	AAA Lys	GAT Asp	TCT Ser	CCA Pro	AAA Lys	GCC Ala	CAG Gln	CTA Leu	CGC Arg	CGG Arg	ACA Thr	ATA Ile	GAA Glu	TGT Cys	TGT Cys	684
110					115					120				125		
CGG Arg	ACC Thr	AAT Asn	TTA Leu	TGT Cys	AAC Asn	CAG Gln	TAT Tyr	TTG Leu	CAA Gln	CCC Pro	ACA Thr	CTG Leu	CCC Pro	CCT Pro	GTT Val	732
				130					135					140		
GTC Val	ATA Ile	GGT Gly	CCG Pro	TTT Phe	TTT Phe	GAT Asp	GGC Gly	AGC Ser	ATT Ile	CGA Arg	TGG Trp	CTG Leu	GTT Val	TTG Leu	CTC Leu	780
		145						150					155			
ATT Ile	TCT Ser	ATG Met	GCT Ala	GTC Val	TGC Cys	ATA Ile	ATT Ile	GCT Ala	ATG Met	ATC Ile	ATC Ile	TTC Phe	TCC Ser	AGC Ser	TGC Cys	828
		160					165					170				
TTT Phe	TGT Cys	TAC Tyr	AAA Lys	CAT His	TAT Tyr	TGC Cys	AAG Lys	AGC Ser	ATC Ile	TCA Ser	AGC Ser	AGA Arg	CGT Arg	CGT Arg	TAC Tyr	876
		175				180					185					
AAT Asn	CGT Arg	GAT Asp	TTG Leu	GAA Glu	CAG Gln	GAT Asp	GAA Glu	GCA Ala	TTT Phe	ATT Ile	CCA Pro	GTT Val	GGA Gly	GAA Glu	TCA Ser	924
190					195					200				205		
CTA Leu	AAA Lys	GAC Asp	CTT Leu	ATT Ile	GAC Asp	CAG Gln	TCA Ser	CAA Gln	AGT Ser	TCT Ser	GGT Gly	AGT Ser	GGG Gly	TCT Ser	GGA Gly	972
				210				215						220		
CTA Leu	CCT Pro	TTA Leu	TTG Leu	GTT Val	CAG Gln	CGA Arg	ACT Thr	ATT Ile	GCC Ala	AAA Lys	CAG Gln	ATT Ile	CAG Gln	ATG Met	GTC Val	1020
			225					230					235			
CGG Arg	CAA Gln	GTT Val	GGT Gly	AAA Lys	GGC Gly	CGA Arg	TAT Tyr	GGA Gly	GAA Glu	GTA Val	TGG Trp	ATG Met	GGC Gly	AAA Lys	TGG Trp	1068
		240					245					250				
CGT Arg	GGC Gly	GAA Glu	AAA Lys	GTG Val	GCG Ala	GTG Val	AAA Lys	GTA Val	TTC Phe	TTT Phe	ACC Thr	ACT Thr	GAA Glu	GAA Glu	GCC Ala	1116
		255				260					265					

AGC Ser 270	TGG Trp	TTT Phe	CGA Arg	GAA Glu	ACA Thr	GAA Glu	ATC Ile	TAC Tyr	CAA Gln	ACT Thr	GTG Val	CTA Leu	ATG Met	CGC Arg	CAT His	1164
275																
GAA Glu	AAC Asn	ATA Ile	CTT Leu	GGT Gly	TTC Phe	ATA Ile	GCG Ala	GCA Ala	GAC Asp	ATT Ile	AAA Lys	GGT Gly	ACA Thr	GGT Gly	TCC Ser	1212
290																
TGG Trp	ACT Thr	CAG Gln	CTC Leu	TAT Tyr	TTG Leu	ATT Ile	ACT Thr	GAT Asp	TAC Tyr	CAT His	GAA Glu	AAT Asn	GGA Gly	TCT Ser	CTC Leu	1260
305																
TAT Tyr	GAC Asp	TTC Phe	CTG Leu	AAA Lys	TGT Cys	GCT Ala	ACA Thr	CTG Leu	GAC Asp	ACC Thr	AGA Arg	GCC Ala	CTG Leu	CTT Leu	AAA Lys	1308
320																
TTG Leu	GCT Ala	TAT Tyr	TCA Ser	GCT Ala	GCC Ala	TGT Cys	GGT Gly	CTG Leu	TGC Cys	CAC His	CTG Leu	CAC His	ACA Thr	GAA Glu	ATT Ile	1356
335																
TAT Tyr	GGC Gly	ACC Thr	CAA Gln	GGA Gly	AAG Lys	CCC Pro	GCA Ala	ATT Ile	GCT Ala	CAT His	CGA Arg	GAC Asp	CTA Leu	AAG Lys	AGC Ser	1404
350																
AAA Lys	AAC Asn	ATC Ile	CTC Leu	ATC Ile	AAG Lys	AAA Lys	AAT Asn	GGG Gly	AGT Ser	TGC Cys	TGC Cys	ATT Ile	GCT Ala	GAC Asp	CTG Leu	1452
370																
GGC Gly	CTT Leu	GCT Ala	GTT Val	AAA Lys	TTC Phe	AAC Asn	AGT Ser	GAC Asp	ACA Thr	AAT Asn	GAA Glu	GTT Val	GAT Asp	GTG Val	CCC Pro	1500
385																
TTG Leu	AAT Asn	ACC Thr	AGG Arg	GTG Val	GGC Gly	ACC Thr	AAA Lys	CGC Arg	TAC Tyr	ATG Met	GCT Ala	CCC Pro	GAA Glu	GTG Val	CTG Leu	1548
400																
GAC Asp	GAA Glu	AGC Ser	CTG Leu	AAC Asn	AAA Lys	AAC Asn	CAC His	TTC Phe	CAG Gln	CCC Pro	TAC Tyr	ATC Ile	ATG Met	GCT Ala	GAC Asp	1596
415																
ATC Ile	TAC Tyr	AGC Ser	TTC Phe	GGC Gly	CTA Leu	ATC Ile	ATT Ile	TGG Trp	GAG Glu	ATG Met	GCT Ala	CGT Arg	CGT Arg	TGT Cys	ATC Ile	1644
430																
ACA Thr	GGA Gly	GGG Gly	ATC Ile	GTG Val	GAA Glu	GAA Glu	TAC Tyr	CAA Gln	TTG Leu	CCA Pro	TAT Tyr	TAC Tyr	AAC Asn	ATG Met	GTA Val	1692
450																
CCG Pro	AGT Ser	GAT Asp	CCG Pro	TCA Ser	TAC Tyr	GAA Glu	GAT Asp	ATG Met	CGT Arg	GAG Glu	GTT Val	GTG Val	TGT Cys	GTC Val	AAA Lys	1740
465																
CGT Arg	TTG Leu	CGG Arg	CCA Pro	ATT Ile	GTG Val	TCT Ser	AAT Asn	CGG Arg	TGG Trp	AAC Asn	AGT Ser	GAT Asp	GAA Glu	TGT Cys	CTA Leu	1788
480																
CGA Arg	GCA Ala	GTT Val	TTG Leu	AAG Lys	CTA Leu	ATG Met	TCA Ser	GAA Glu	TGC Cys	TGG Trp	GCC Ala	CAC His	AAT Asn	CCA Pro	GCC Ala	1836
495																
500																
505																

TCC AGA CTC ACA GCA TTG AGA ATT AAG AAG ACG CTT GCC AAG ATG GTT	1884
Ser Arg Leu Thr Ala Leu Arg Ile Lys Lys Thr Leu Ala Lys Met Val	
510 515 520 525	
GAA TCC CAA GAT GTA AAA ATC TGATGGTTAA ACCATCGGAG GAGAAACTCT	1935
Glu Ser Gln Asp Val Lys Ile	
530	
AGACTGCAAG AACTGTTTTT ACCCATGGCA TGGGTGGAAT TAGAGTGGAA TAAGGATGTT	1995
AACTTGGTTC TCAGACTCTT TCTTCACTAC GTGTTCACAG GCTGCTAATA TTAACCTTT	2055
CAGTACTCTT ATTAGGATAC AAGCTGGGAA CTTCTAAACA CTTCATTCTT TATATATGGA	2115
CAGCTTTATT TTAATGTGG TTTTGTATGC CTTTTTTTAA GTGGGTTTTT ATGAACTGCA	2175
TCAAGACTTC AATCCTGATT AGTGTCTCCA GTCAGCTCT GGGTACTGAA TTGCCTGTTC	2235
ATAAAACGGT GCTTCTGTG AAAGCCTTAA GAAGATAAAT GAGCCGAGCA GAGATGGAGA	2295
AATAGACTTT GCCTTTTACC TGAGACATTC AGTTCGTTG TATTCTACCT TTGTAAAACA	2355
GCCTATAGAT GATGATGTGT TTGGGATACT GCTTATTTA TGATAGTTG TCCTGTGTCC	2415
TTAGTGATGT GTGTGTGTCT CCATGCACAT GCACGCCGGG ATTCCTCTGC TGCCATTGA	2475
ATTAGAAGAA AATAATTTAT ATGCATGCAC AGGAAGATAT TGGTGGCCGG TGGTTTTGTG	2535
CTTTAAAAT GCAATATCTG ACCAAGATTC GCCAATCTCA TACAAGCCAT TTACTTTGCA	2595
AGTGAGATAG CTTCCCCACC AGCTTTATTT TTTAACATGA AAGCTGATGC CAAGGCCAAA	2655
AGAAGTTTAA AGCATCTGTA AATTGGACT GTTTTCCTTC AACCACCATT TTTTTGTGG	2715
TTATTATTTT TGTCACGGAA AGCATCCTCT CCAAGTTGG AGCTTCTATT GCCATGAACC	2775
ATGCTTACAA AGAAGCACT TCTTATGAA GTGAATTCCT GCATTTGATA GCAATGTAAG	2835
TGCCTATAAC CATGTTCTAT ATTCTTTATT CTCAGTAACT TTTAAAAGGG AAGTTATTTA	2895
TATTTTGTGT ATAATGTGCT TTATTTGCAA ATCACCC	2932

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 532 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met	Thr	Gln	Leu	Tyr	Ile	Tyr	Ile	Arg	Leu	Leu	Gly	Ala	Tyr	Leu	Phe
1				5					10					15	
Ile	Ile	Ser	Arg	Val	Gln	Gly	Gln	Asn	Leu	Asp	Ser	Met	Leu	His	Gly
				20				25					30		

Thr Gly Met Lys Ser Asp Ser Asp Gln Lys Lys Ser Glu Asn Gly Val
 35 40 45
 Thr Leu Ala Pro Glu Asp Thr Leu Pro Phe Leu Lys Cys Tyr Cys Ser
 50 55 60
 Gly His Cys Pro Asp Asp Ala Ile Asn Asn Thr Cys Ile Thr Asn Gly
 65 70 75 80
 His Cys Phe Ala Ile Ile Glu Glu Asp Asp Gln Gly Glu Thr Thr Leu
 85 90 95
 Ala Ser Gly Cys Met Lys Tyr Glu Gly Ser Asp Phe Gln Cys Lys Asp
 100 105 110
 Ser Pro Lys Ala Gln Leu Arg Arg Thr Ile Glu Cys Cys Arg Thr Asn
 115 120 125
 Leu Cys Asn Gln Tyr Leu Gln Pro Thr Leu Pro Pro Val Val Ile Gly
 130 135 140
 Pro Phe Phe Asp Gly Ser Ile Arg Trp Leu Val Leu Leu Ile Ser Met
 145 150 155 160
 Ala Val Cys Ile Ile Ala Met Ile Ile Phe Ser Ser Cys Phe Cys Tyr
 165 170 175
 Lys His Tyr Cys Lys Ser Ile Ser Ser Arg Arg Arg Tyr Asn Arg Asp
 180 185 190
 Leu Glu Gln Asp Glu Ala Phe Ile Pro Val Gly Glu Ser Leu Lys Asp
 195 200 205
 Leu Ile Asp Gln Ser Gln Ser Ser Gly Ser Gly Ser Gly Leu Pro Leu
 210 215 220
 Leu Val Gln Arg Thr Ile Ala Lys Gln Ile Gln Met Val Arg Gln Val
 225 230 235 240
 Gly Lys Gly Arg Tyr Gly Glu Val Trp Met Gly Lys Trp Arg Gly Glu
 245 250 255
 Lys Val Ala Val Lys Val Phe Phe Thr Thr Glu Glu Ala Ser Trp Phe
 260 265 270
 Arg Glu Thr Glu Ile Tyr Gln Thr Val Leu Met Arg His Glu Asn Ile
 275 280 285
 Leu Gly Phe Ile Ala Ala Asp Ile Lys Gly Thr Gly Ser Trp Thr Gln
 290 295 300
 Leu Tyr Leu Ile Thr Asp Tyr His Glu Asn Gly Ser Leu Tyr Asp Phe
 305 310 315 320
 Leu Lys Cys Ala Thr Leu Asp Thr Arg Ala Leu Leu Lys Leu Ala Tyr
 325 330 335
 Ser Ala Ala Cys Gly Leu Cys His Leu His Thr Glu Ile Tyr Gly Thr
 340 345 350

Gln Gly Lys Pro Ala Ile Ala His Arg Asp Leu Lys Ser Lys Asn Ile
 355 360 365
 Leu Ile Lys Lys Asn Gly Ser Cys Cys Ile Ala Asp Leu Gly Leu Ala
 370 375 380
 Val Lys Phe Asn Ser Asp Thr Asn Glu Val Asp Val Pro Leu Asn Thr
 385 390 395 400
 Arg Val Gly Thr Lys Arg Tyr Met Ala Pro Glu Val Leu Asp Glu Ser
 405 410 415
 Leu Asn Lys Asn His Phe Gln Pro Tyr Ile Met Ala Asp Ile Tyr Ser
 420 425 430
 Phe Gly Leu Ile Ile Trp Glu Met Ala Arg Arg Cys Ile Thr Gly Gly
 435 440 445
 Ile Val Glu Glu Tyr Gln Leu Pro Tyr Tyr Asn Met Val Pro Ser Asp
 450 455 460
 Pro Ser Tyr Glu Asp Met Arg Glu Val Val Cys Val Lys Arg Leu Arg
 465 470 475 480
 Pro Ile Val Ser Asn Arg Trp Asn Ser Asp Glu Cys Leu Arg Ala Val
 485 490 495
 Leu Lys Leu Met Ser Glu Cys Trp Ala His Asn Pro Ala Ser Arg Leu
 500 505 510
 Thr Ala Leu Arg Ile Lys Lys Thr Leu Ala Lys Met Val Glu Ser Gln
 515 520 525
 Asp Val Lys Ile
 530

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2333 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (v) FRAGMENT TYPE: internal
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 1..1515

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

ATG GCG GAG TCG GCC GGA GCC TCC TCC TTC TTC CCC CTT GTT GTC CTC	48
Met Ala Glu Ser Ala Gly Ala Ser Ser Phe Phe Pro Leu Val Val Leu	
1 5 10 15	
CTG CTC GCC GGC AGC GGC GGG TCC GGG CCC CGG GGG GTC CAG GCT CTG	96
Leu Leu Ala Gly Ser Gly Gly Ser Gly Pro Arg Gly Val Gln Ala Leu	
20 25 30	
CTG TGT GCG TGC ACC AGC TGC CTC CAG GCC AAC TAC ACG TGT GAG ACA	144
Leu Cys Ala Cys Thr Ser Cys Leu Gln Ala Asn Tyr Thr Cys Glu Thr	
35 40 45	
GAT GGG GCC TGC ATG GTT TCC TTT TTC AAT CTG GAT GGG ATG GAG CAC	192
Asp Gly Ala Cys Met Val Ser Phe Phe Asn Leu Asp Gly Met Glu His	
50 55 60	
CAT GTG CGC ACC TGC ATC CCC AAA GTG GAG CTG GTC CCT GCC GGG AAG	240
His Val Arg Thr Cys Ile Pro Lys Val Glu Leu Val Pro Ala Gly Lys	
65 70 75 80	
CCC TTC TAC TGC CTG AGC TCG GAG GAC CTG CGC AAC ACC CAC TGC TGC	288
Pro Phe Tyr Cys Leu Ser Ser Glu Asp Leu Arg Asn Thr His Cys Cys	
85 90 95	
TAC ACT GAC TAC TGC AAC AGG ATC GAC TTG AGG GTG CCC AGT GGT CAC	336
Tyr Thr Asp Tyr Cys Asn Arg Ile Asp Leu Arg Val Pro Ser Gly His	
100 105 110	
CTC AAG GAG CCT GAG CAC CCG TCC ATG TGG GGC CCG GTG GAG CTG GTA	384
Leu Lys Glu Pro Glu His Pro Ser Met Trp Gly Pro Val Glu Leu Val	
115 120 125	
GGC ATC ATC GCC GGC CCG GTG TTC CTC CTG TTC CTC ATC ATC ATC ATT	432
Gly Ile Ile Ala Gly Pro Val Phe Leu Leu Phe Leu Ile Ile Ile Ile	
130 135 140	
GTT TTC CTT GTC ATT AAC TAT CAT CAG CGT GTC TAT CAC AAC CGC CAG	480
Val Phe Leu Val Ile Asn Tyr His Gln Arg Val Tyr His Asn Arg Gln	
145 150 155 160	
AGA CTG GAC ATG GAA GAT CCC TCA TGT GAG ATG TGT CTC TCC AAA GAC	528
Arg Leu Asp Met Glu Asp Pro Ser Cys Glu Met Cys Leu Ser Lys Asp	
165 170 175	
AAG ACG CTC CAG GAT CTT GTC TAC GAT CTC TCC ACC TCA GGG TCT GGC	576
Lys Thr Leu Gln Asp Leu Val Tyr Asp Leu Ser Thr Ser Gly Ser Gly	
180 185 190	
TCA GGG TTA CCC CTC TTT GTC CAG CGC ACA GTG GCC CGA ACC ATC GTT	624
Ser Gly Leu Pro Leu Phe Val Gln Arg Thr Val Ala Arg Thr Ile Val	
195 200 205	
TTA CAA GAG ATT ATT GGC AAG GGT CGG TTT GGG GAA GTA TGG CGG GGC	672
Leu Gln Glu Ile Ile Gly Lys Gly Arg Phe Gly Glu Val Trp Arg Gly	
210 215 220	

CGC Arg 225	TGG Trp	AGG Arg	GCT Gly	GGT Gly	GAT Asp	GTG Val	GCT Ala	GTG Val	AAA Lys	ATA Ile	TTC Phe	TCT Ser	TCT Ser	CGT Arg	GAA Glu	240	720
GAA Glu	CGG Arg	TCT Ser	TGG Trp	TTC Phe	AGG Arg	GAA Glu	GCA Ala	GAG Glu	ATA Ile	TAC Tyr	CAG Gln	ACG Thr	GTC Val	ATG Met	CTG Leu	255	768
CGC Arg	CAT His	GAA Glu	AAC Asn	ATC Ile	CTT Leu	GGA Gly	TTT Phe	ATT Ile	GCT Ala	GCT Ala	GAC Asp	AAT Asn	AAA Lys	GAT Asp	AAT Asn	270	816
GGC Gly	ACC Thr	TGG Trp	ACA Thr	CAG Gln	CTG Leu	TGG Trp	CTT Leu	GTT Val	TCT Ser	GAC Asp	TAT Tyr	CAT His	GAG Glu	CAC His	GGG Gly	285	864
TCC Ser	CTG Leu	TTT Phe	GAT Asp	TAT Tyr	CTG Leu	AAC Asn	CGG Arg	TAC Tyr	ACA Thr	GTG Val	ACA Thr	ATT Ile	GAG Glu	GGG Gly	ATG Met	300	912
ATT Ile	AAG Lys	CTG Leu	GCC Ala	TTG Leu	TCT Ser	GCT Ala	GCT Ala	AGT Ser	GGG Gly	CTG Leu	GCA Ala	CAC His	CTG Leu	CAC His	ATG Met	320	960
GAG Glu	ATC Ile	GTG Val	GGC Gly	ACC Thr	CAA Gln	GGG Gly	AAG Lys	CCT Pro	GGA Gly	ATT Ile	GCT Ala	CAT His	CGA Arg	GAC Asp	TTA Leu	335	1008
AAG Lys	TCA Ser	AAG Lys	AAC Asn	ATT Ile	CTG Leu	GTG Val	AAG Lys	AAA Lys	AAT Asn	GGC Gly	ATG Met	TGT Cys	GCC Ala	ATA Ile	GCA Ala	350	1056
GAC Asp	CTG Leu	GGC Gly	CTG Leu	GCT Ala	GTC Val	CGT Arg	CAT His	GAT Asp	GCA Ala	GTC Val	ACT Thr	GAC Asp	ACC Thr	ATT Ile	GAC Asp	365	1104
ATT Ile	GCC Ala	CCG Pro	AAT Asn	CAG Gln	AGG Arg	GTG Val	GGG Gly	ACC Thr	AAA Lys	CGA Arg	TAC Tyr	ATG Met	GCC Ala	CCT Pro	GAA Glu	380	1152
GTA Val	CTT Leu	GAT Asp	GAA Glu	ACC Thr	ATT Ile	AAT Asn	ATG Met	AAA Lys	CAC His	TTT Phe	GAC Asp	TCC Ser	TTT Phe	AAA Lys	TGT Cys	400	1200
GCT Ala	GAT Asp	ATT Ile	TAT Tyr	GCC Ala	CTC Leu	GGG Gly	CTT Leu	GTA Val	TAT Tyr	TGG Trp	GAG Glu	ATT Ile	GCT Ala	CGA Arg	AGA Arg	415	1248
TGC Cys	AAT Asn	TCT Ser	GGA Gly	GGA Gly	GTC Val	CAT His	GAA Glu	GAA Glu	TAT Tyr	CAG Gln	CTG Leu	CCA Pro	TAT Tyr	TAC Tyr	GAC Asp	430	1296
TTA Leu	GTG Val	CCC Pro	TCT Ser	GAC Asp	CCT Pro	TCC Ser	ATT Ile	GAG Glu	GAA Glu	ATG Met	CGA Arg	AAG Lys	GTT Val	GTA Val	TGT Cys	445	1344
GAT Asp	CAG Gln	AAG Lys	CTG Leu	CGT Arg	CCC Pro	AAC Asn	ATC Ile	CCC Pro	AAC Asn	TGG Trp	TGG Trp	CAG Gln	AGT Ser	TAT Tyr	GAG Glu	460	1392

GCA CTG CGG GTG ATG GGG AAG ATG ATG CGA GAG TGT TGG TAT GCC AAC	1440
Ala Leu Arg Val Met Gly Lys Met Met Arg Glu Cys Trp Tyr Ala Asn	
465 470 475 480	
GGC GCA GCC CGC CTG ACG GCC CTG CGC ATC AAG AAG ACC CTC TCC CAG	1488
Gly Ala Ala Arg Leu Thr Ala Leu Arg Ile Lys Lys Thr Leu Ser Gln	
485 490 495	
CTC AGC GTG CAG GAA GAC GTG AAG ATC TAACTGCTCC CTCTCTCCAC	1535
Leu Ser Val Gln Glu Asp Val Lys Ile	
500 505	
ACGGAGCTCC TGGCAGCGAG AACTAGGCAC AGCTGCCGCG TTGAGCGTAC GATGGAGGCC	1595
TACCTCTCGT TTCTGCCAG CCCTCTGTGG CCAGGAGCCC TGGCCCGCAA GAGGGACAGA	1655
GCCCCGGAGA GACTCGCTCA CTCCCATGTT GGGTTTGAGA CAGACACCTT TTCTATTTAC	1715
CTCCTAATGG CATGGAGACT CTGAGAGCGA ATTGTGTGGA GAACTCAGTG CCACACCTCG	1775
AACTGGTTGT AGTGGGAAGT CCGCGAAAC CCGGTGCATC TGGCACGTGG CCAGGAGCCA	1835
TGACAGGGGC GCTTGGGAGG GGCCGGAGGA ACCGAGGTGT TGCCAGTGCT AAGCTGCCCT	1895
GAGGGTTTCC TTCGGGGACC AGCCACAGC ACACCAAGGT GGCCCGGAAG AACCAGAAGT	1955
GCAGCCCCTC TCACAGGCAG CTCTGAGCCG CGTTTCCCC TCCTCCCTGG GATGGAGCCT	2015
GCCGGGAGAC TGCCAGTGA GACGGAATCT GCCGTTTGT CTGTCCAGCC GTGTGTGCAT	2075
GTGCCGAGGT GCCTCCCCCG TTGTGCCTGG TTCGTGCCAT GCCCTTACAC GTGCGTGTGA	2135
GTGTGTGTGT GTGTCTGTAG GTGCGCACTT ACCTGCTTGA GCTTCTGTG CATGTGCAGG	2195
TGGGGGTGT GGTGTCATG CTGTCCGTGC TTGCTGGTGC CTCTTTCAG TAGTGACCAG	2255
CATCTAGTTT CCCTGGTGCC CTTCCTGGA GGTCTCTCCC TCCCCAGAG CCCCTCATGC	2315
CACAGTGGTA CTCTGTGT	2333

(2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 505 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Met	Ala	Glu	Ser	Ala	Gly	Ala	Ser	Ser	Phe	Phe	Pro	Leu	Val	Val	Leu
1				5					10					15	
Leu	Leu	Ala	Gly	Ser	Gly	Gly	Ser	Gly	Pro	Arg	Gly	Val	Gln	Ala	Leu
			20					25					30		

Leu Cys Ala Cys Thr Ser Cys Leu Gln Ala Asn Tyr Thr Cys Glu Thr
 35 40 45
 Asp Gly Ala Cys Met Val Ser Phe Phe Asn Leu Asp Gly Met Glu His
 50 55 60
 His Val Arg Thr Cys Ile Pro Lys Val Glu Leu Val Pro Ala Gly Lys
 65 70 75 80
 Pro Phe Tyr Cys Leu Ser Ser Glu Asp Leu Arg Asn Thr His Cys Cys
 85 90 95
 Tyr Thr Asp Tyr Cys Asn Arg Ile Asp Leu Arg Val Pro Ser Gly His
 100 105 110
 Leu Lys Glu Pro Glu His Pro Ser Met Trp Gly Pro Val Glu Leu Val
 115 120 125
 Gly Ile Ile Ala Gly Pro Val Phe Leu Leu Phe Leu Ile Ile Ile Ile
 130 135 140
 Val Phe Leu Val Ile Asn Tyr His Gln Arg Val Tyr His Asn Arg Gln
 145 150 155 160
 Arg Leu Asp Met Glu Asp Pro Ser Cys Glu Met Cys Leu Ser Lys Asp
 165 170 175
 Lys Thr Leu Gln Asp Leu Val Tyr Asp Leu Ser Thr Ser Gly Ser Gly
 180 185 190
 Ser Gly Leu Pro Leu Phe Val Gln Arg Thr Val Ala Arg Thr Ile Val
 195 200 205
 Leu Gln Glu Ile Ile Gly Lys Gly Arg Phe Gly Glu Val Trp Arg Gly
 210 215 220
 Arg Trp Arg Gly Gly Asp Val Ala Val Lys Ile Phe Ser Ser Arg Glu
 225 230 235 240
 Glu Arg Ser Trp Phe Arg Glu Ala Glu Ile Tyr Gln Thr Val Met Leu
 245 250 255
 Arg His Glu Asn Ile Leu Gly Phe Ile Ala Ala Asp Asn Lys Asp Asn
 260 265 270
 Gly Thr Trp Thr Gln Leu Trp Leu Val Ser Asp Tyr His Glu His Gly
 275 280 285
 Ser Leu Phe Asp Tyr Leu Asn Arg Tyr Thr Val Thr Ile Glu Gly Met
 290 295 300
 Ile Lys Leu Ala Leu Ser Ala Ala Ser Gly Leu Ala His Leu His Met
 305 310 315 320
 Glu Ile Val Gly Thr Gln Gly Lys Pro Gly Ile Ala His Arg Asp Leu
 325 330 335
 Lys Ser Lys Asn Ile Leu Val Lys Lys Asn Gly Met Cys Ala Ile Ala
 340 345 350

Asp Leu Gly Leu Ala Val Arg His Asp Ala Val Thr Asp Thr Ile Asp
 355 360 365

Ile Ala Pro Asn Gln Arg Val Gly Thr Lys Arg Tyr Met Ala Pro Glu
 370 375 380

Val Leu Asp Glu Thr Ile Asn Met Lys His Phe Asp Ser Phe Lys Cys
 385 390 395 400

Ala Asp Ile Tyr Ala Leu Gly Leu Val Tyr Trp Glu Ile Ala Arg Arg
 405 410 415

Cys Asn Ser Gly Gly Val His Glu Glu Tyr Gln Leu Pro Tyr Tyr Asp
 420 425 430

Leu Val Pro Ser Asp Pro Ser Ile Glu Glu Met Arg Lys Val Val Cys
 435 440 445

Asp Gln Lys Leu Arg Pro Asn Ile Pro Asn Trp Trp Gln Ser Tyr Glu
 450 455 460

Ala Leu Arg Val Met Gly Lys Met Met Arg Glu Cys Trp Tyr Ala Asn
 465 470 475 480

Gly Ala Ala Arg Leu Thr Ala Leu Arg Ile Lys Lys Thr Leu Ser Gln
 485 490 495

Leu Ser Val Gln Glu Asp Val Lys Ile
 500 505

(2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2308 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Mouse

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 77..1585

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GGCGAGGCGA GGTTTGCTGG GGTGAGGCAG CGGCGCGGCC GGGCCGGGCC GGGCCACAGG

60

CGGTGGCGGC	GGGACC	ATG	GAG	GCG	GCG	GTC	GCT	GCT	CCG	CGT	CCC	CGG		109		
		Met	Glu	Ala	Ala	Val	Ala	Ala	Pro	Arg	Pro	Arg				
		1				5					10					
CTG	CTC	CTC	CTC	GTG	CTG	GCG	GCG	GCG	GCG	GCG	GCG	GCG	CTG	157		
Leu	Leu	Leu	Leu	Val	Leu	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Leu		
			15				20					25				
CTC	CCG	GGG	GCG	ACG	GCG	TTA	CAG	TGT	TTC	TGC	CAC	CTC	TGT	ACA	AAA	205
Leu	Pro	Gly	Ala	Thr	Ala	Leu	Gln	Cys	Phe	Cys	His	Leu	Cys	Thr	Lys	
		30					35					40				
GAC	AAT	TTT	ACT	TGT	GTG	ACA	GAT	GGG	CTC	TGC	TTT	GTC	TCT	GTC	ACA	253
Asp	Asn	Phe	Thr	Cys	Val	Thr	Asp	Gly	Leu	Cys	Phe	Val	Ser	Val	Thr	
	45					50					55					
GAG	ACC	ACA	GAC	AAA	GTT	ATA	CAC	AAC	AGC	ATG	TGT	ATA	GCT	GAA	ATT	301
Glu	Thr	Thr	Asp	Lys	Val	Ile	His	Asn	Ser	Met	Cys	Ile	Ala	Glu	Ile	
	60				65					70					75	
GAC	TTA	ATT	CCT	CGA	GAT	AGG	CCG	TTT	GTA	TGT	GCA	CCC	TCT	TCA	AAA	349
Asp	Leu	Ile	Pro	Arg	Asp	Arg	Pro	Phe	Val	Cys	Ala	Pro	Ser	Ser	Lys	
				80					85						90	
ACT	GGG	TCT	GTG	ACT	ACA	ACA	TAT	TGC	TGC	AAT	CAG	GAC	CAT	TGC	AAT	397
Thr	Gly	Ser	Val	Thr	Thr	Thr	Tyr	Cys	Cys	Asn	Gln	Asp	His	Cys	Asn	
			95					100					105			
AAA	ATA	GAA	CTT	CCA	ACT	ACT	GTA	AAG	TCA	TCA	CCT	GGC	CTT	GGT	CCT	445
Lys	Ile	Glu	Leu	Pro	Thr	Thr	Val	Lys	Ser	Ser	Pro	Gly	Leu	Gly	Pro	
		110					115					120				
GTG	GAA	CTG	GCA	GCT	GTC	ATT	GCT	GGA	CCA	GTG	TGC	TTC	GTC	TGC	ATC	493
Val	Glu	Leu	Ala	Ala	Val	Ile	Ala	Gly	Pro	Val	Cys	Phe	Val	Cys	Ile	
	125					130					135					
TCA	CTC	ATG	TTG	ATG	GTC	TAT	ATC	TGC	CAC	AAC	CGC	ACT	GTC	ATT	CAC	541
Ser	Leu	Met	Leu	Met	Val	Tyr	Ile	Cys	His	Asn	Arg	Thr	Val	Ile	His	
					145					150					155	
CAT	CGA	GTG	CCA	AAT	GAA	GAG	GAC	CCT	TCA	TTA	GAT	CGC	CCT	TTT	ATT	589
His	Arg	Val	Pro	Asn	Glu	Glu	Asp	Pro	Ser	Leu	Asp	Arg	Pro	Phe	Ile	
				160					165					170		
TCA	GAG	GGT	ACT	ACG	TTG	AAA	GAC	TTA	ATT	TAT	GAT	ATG	ACA	ACG	TCA	637
Ser	Glu	Gly	Thr	Thr	Leu	Lys	Asp	Leu	Ile	Tyr	Asp	Met	Thr	Thr	Ser	
			175					180					185			
GGT	TCT	GGC	TCA	GGT	TTA	CCA	TTG	CTT	GTT	CAG	AGA	ACA	ATT	GCG	AGA	685
Gly	Ser	Gly	Ser	Gly	Leu	Pro	Leu	Leu	Val	Gln	Arg	Thr	Ile	Ala	Arg	
		190					195					200				
ACT	ATT	GTG	TTA	CAA	GAA	AGC	ATT	GGC	AAA	GGT	CGA	TTT	GGA	GAA	GTT	733
Thr	Ile	Val	Leu	Gln	Glu	Ser	Ile	Gly	Lys	Gly	Arg	Phe	Gly	Glu	Val	
	205					210					215					
TGC	AGA	GGA	AAG	TGG	CCG	GGA	GAA	GAA	GTT	GCT	GTT	AAG	ATA	TTC	TCC	781
Trp	Arg	Gly	Lys	Trp	Arg	Gly	Glu	Glu	Val	Ala	Val	Lys	Ile	Phe	Ser	
					225					230					235	

TCT Ser	AGA Arg	GAA Glu	GAA Glu	CGT Arg	TCG Ser	TGG Trp	TTC Phe	CGT Arg	GAG Glu	GCA Ala	GAG Glu	ATT Ile	TAT Tyr	CAA Gln	ACT Thr	829
				240				245						250		
GTA Val	ATG Met	TTA Leu	CGT Arg	CAT His	GAA Glu	AAC Asn	ATC Ile	CTG Leu	GGA Gly	TTT Phe	ATA Ile	GCA Ala	GCA Ala	GAC Asp	AAT Asn	877
			255					260						265		
AAA Lys	GAC Asp	AAT Asn	GGT Gly	ACT Thr	TGG Trp	ACT Thr	CAG Gln	CTC Leu	TGG Trp	TTG Leu	GTG Val	TCA Ser	GAT Asp	TAT Tyr	CAT His	925
			270				275							280		
GAG Glu	CAT His	GGA Gly	TCC Ser	CTT Leu	TTT Phe	GAT Asp	TAC Tyr	TTA Leu	AAC Asn	AGA Arg	TAC Tyr	ACA Thr	GTT Val	ACT Thr	GTG Val	973
			285			290										
GAA Glu	GGA Gly	ATG Met	ATA Ile	AAA Lys	CTT Leu	GCT Ala	CTG Leu	TCC Ser	ACG Thr	GCG Ala	AGC Ser	GGT Gly	CTT Leu	GCC Ala	CAT His	1021
					300	305				310					315	
CTT Leu	CAC His	ATG Met	GAG Glu	ATT Ile	GTT Val	GGT Gly	ACC Thr	CAA Gln	GGA Gly	AAG Lys	CCA Pro	GCC Ala	ATT Ile	GCT Ala	CAT His	1069
					320				325						330	
AGA Arg	GAT Asp	TTG Leu	AAA Lys	TCA Ser	AAG Lys	AAT Asn	ATC Ile	TTG Leu	GTA Val	AAG Lys	AAG Lys	AAT Asn	GGA Gly	ACT Thr	TGC Cys	1117
					335			340						345		
TGT Cys	ATT Ile	GCA Ala	GAC Asp	TTA Leu	GGA Gly	CTG Leu	GCA Ala	GTA Val	AGA Arg	CAT His	GAT Asp	TCA Ser	GCC Ala	ACA Thr	GAT Asp	1165
			350			355								360		
ACC Thr	ATT Ile	GAT Asp	ATT Ile	GCT Ala	CCA Pro	AAC Asn	CAC His	AGA Arg	GTG Val	GGA Gly	ACA Thr	AAA Lys	AGG Arg	TAC Tyr	ATG Met	1213
					365		370							375		
GCC Ala	CCT Pro	GAA Glu	GTT Val	CTC Leu	GAT Asp	GAT Asp	TCC Ser	ATA Ile	AAT Asn	ATG Met	AAA Lys	CAT His	TTT Phe	GAA Glu	TCC Ser	1261
					380		385			390					395	
TTC Phe	AAA Lys	CGT Arg	GCT Ala	GAC Asp	ATC Ile	TAT Tyr	GCA Ala	ATG Met	GGC Gly	TTA Leu	GTA Val	TTC Phe	TGG Trp	GAA Glu	ATT Ile	1309
				400					405					410		
GCT Ala	CGA Arg	CGA Arg	TGT Cys	TCC Ser	ATT Ile	GGT Gly	GGA Gly	ATT Ile	CAT His	GAA Glu	GAT Asp	TAC Tyr	CAA Gln	CTG Leu	CCT Pro	1357
			415					420						425		
TAT Tyr	TAT Tyr	GAT Asp	CTT Leu	GTA Val	CCT Pro	TCT Ser	GAC Asp	CCA Pro	TCA Ser	GTT Val	GAA Glu	GAA Glu	ATG Met	AGA Arg	AAA Lys	1405
			430				435						440			
GTT Val	GTT Val	TGT Cys	GAA Glu	CAG Gln	AAG Lys	TTA Leu	AGG Arg	CCA Pro	AAT Asn	ATC Ile	CCA Pro	AAC Asn	AGA Arg	TGG Trp	CAG Gln	1453
			445				450							455		
AGC Ser	TGT Cys	GAA Glu	GCC Ala	TTG Leu	AGA Arg	GTA Val	ATG Met	GCT Ala	AAA Lys	ATT Ile	ATG Met	AGA Arg	GAA Glu	TGT Cys	TGG Trp	1501
					460		465				470				475	

TAT GCC AAT GGA GCA GCT AGG CTT ACA GCA TTG CGG ATT AAG AAA ACA	1549
Tyr Ala Asn Gly Ala Ala Arg Leu Thr Ala Leu Arg Ile Lys Lys Thr	
480 485 490	
TTA TCG CAA CTC AGT CAA CAG GAA GGC ATC AAA ATG TAATTCTACA	1595
Leu Ser Gln Leu Ser Gln Gln Glu Gly Ile Lys Met	
495 500	
GCTTTGCCCTG AACTCTCCTT TTTTCTTCAG ATCTGCTCCT GGGTTTTAAT TTGGGAGGTC	1655
AGTTGTCTA CCTCACTGAG AGGGAACAGA AGGATATTGC TTCCTTTTGC AGCAGTGTA	1715
TAAAGTCAAT TAAAACTTC CCAGGATTC TTTGGACCCA GGAAACAGCC ATGTGGGTCC	1775
TTTCTGTGCA CTATGAACGC TTCTTTCCCA GGACAGAAAA TGTGTAGTCT ACCTTTATTT	1835
TTTATTAACA AAACCTGTTT TTTAAAAAGA TGATTGCTGG TCTTAACTTT AGGTAACCTCT	1895
GCTGTGCTGG AGATCATCTT TAAGGGCAA GGAGTTGGAT TGCTGAATTA CAATGAAACA	1955
TGTCTTATTA CTAAGAAAG TGATTTACTC CTGGTTAGTA CATTCTCAGA GGATTCTGAA	2015
CCACTAGAGT TTCCTTGATT CAGACTTTGA ATGTAAGTGT CTATAGTTTT TCAGGATCTT	2075
AAAACATAACA CTTATAAAC TCTTATCTTG AGTCTAAAA TGACCTCATA TAGTAGTGAG	2135
GAACATAATT CATGCAATTG TATTTTGTAT ACTATTATTG TTCTTTCACT TATTCAGAAC	2195
ATTACATGCC TTCAAATGG GATTGTACTA TACCAGTAAG TGCCACTTCT GTGTCTTTCT	2255
AATGGAATG AGTAGAATTG CTGAAAGTCT CTATGTTAAA ACCTATAGTG TTT	2308

(2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 503 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Met	Glu	Ala	Ala	Val	Ala	Ala	Pro	Arg	Pro	Arg	Leu	Leu	Leu	Leu	Val
1				5				10						15	
Leu	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Leu	Leu	Pro	Gly	Ala	Thr
			20					25					30		
Ala	Leu	Gln	Cys	Phe	Cys	His	Leu	Cys	Thr	Lys	Asp	Asn	Phe	Thr	Cys
		35					40					45			
Val	Thr	Asp	Gly	Leu	Cys	Phe	Val	Ser	Val	Thr	Glu	Thr	Thr	Asp	Lys
	50					55					60				
Val	Ile	His	Asn	Ser	Met	Cys	Ile	Ala	Glu	Ile	Asp	Leu	Ile	Pro	Arg
65					70					75					80

Asp Arg Pro Phe Val Cys Ala Pro Ser Ser Lys Thr Gly Ser Val Thr
 85 90 95
 Thr Thr Tyr Cys Cys Asn Gln Asp His Cys Asn Lys Ile Glu Leu Pro
 100 105 110
 Thr Thr Val Lys Ser Ser Pro Gly Leu Gly Pro Val Glu Leu Ala Ala
 115 120 125
 Val Ile Ala Gly Pro Val Cys Phe Val Cys Ile Ser Leu Met Leu Met
 130 135 140
 Val Tyr Ile Cys His Asn Arg Thr Val Ile His His Arg Val Pro Asn
 145 150 155 160
 Glu Glu Asp Pro Ser Leu Asp Arg Pro Phe Ile Ser Glu Gly Thr Thr
 165 170 175
 Leu Lys Asp Leu Ile Tyr Asp Met Thr Thr Ser Gly Ser Gly Ser Gly
 180 185 190
 Leu Pro Leu Leu Val Gln Arg Thr Ile Ala Arg Thr Ile Val Leu Gln
 195 200 205
 Glu Ser Ile Gly Lys Gly Arg Phe Gly Glu Val Trp Arg Gly Lys Trp
 210 215 220
 Arg Gly Glu Glu Val Ala Val Lys Ile Phe Ser Ser Arg Glu Glu Arg
 225 230 235 240
 Ser Trp Phe Arg Glu Ala Glu Ile Tyr Gln Thr Val Met Leu Arg His
 245 250 255
 Glu Asn Ile Leu Gly Phe Ile Ala Ala Asp Asn Lys Asp Asn Gly Thr
 260 265 270
 Trp Thr Gln Leu Trp Leu Val Ser Asp Tyr His Glu His Gly Ser Leu
 275 280 285
 Phe Asp Tyr Leu Asn Arg Tyr Thr Val Thr Val Glu Gly Met Ile Lys
 290 295 300
 Leu Ala Leu Ser Thr Ala Ser Gly Leu Ala His Leu His Met Glu Ile
 305 310 315 320
 Val Gly Thr Gln Gly Lys Pro Ala Ile Ala His Arg Asp Leu Lys Ser
 325 330 335
 Lys Asn Ile Leu Val Lys Lys Asn Gly Thr Cys Cys Ile Ala Asp Leu
 340 345 350
 Gly Leu Ala Val Arg His Asp Ser Ala Thr Asp Thr Ile Asp Ile Ala
 355 360 365
 Pro Asn His Arg Val Gly Thr Lys Arg Tyr Met Ala Pro Glu Val Leu
 370 375 380
 Asp Asp Ser Ile Asn Met Lys His Phe Glu Ser Phe Lys Arg Ala Asp
 385 390 395 400

Ile Tyr Ala Met Gly Leu Val Phe Trp Glu Ile Ala Arg Arg Cys Ser
 405 410 415
 Ile Gly Gly Ile His Glu Asp Tyr Gln Leu Pro Tyr Tyr Asp Leu Val
 420 425 430
 Pro Ser Asp Pro Ser Val Glu Glu Met Arg Lys Val Val Cys Glu Gln
 435 440 445
 Lys Leu Arg Pro Asn Ile Pro Asn Arg Trp Gln Ser Cys Glu Ala Leu
 450 455 460
 Arg Val Met Ala Lys Ile Met Arg Glu Cys Trp Tyr Ala Asn Gly Ala
 465 470 475 480
 Ala Arg Leu Thr Ala Leu Arg Ile Lys Lys Thr Leu Ser Gln Leu Ser
 485 490 495
 Gln Gln Glu Gly Ile Lys Met
 500

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1922 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (v) FRAGMENT TYPE: internal
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Mouse
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 241..1746

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GAGAGCACAG CCCTTCCCAG TCCCCGGAGC CGCCGCGCCA CGCCGCGCATG ATCAAGACCT	60
TTTCCCCGGC CCCACAGGGC CTCTGGACGT GAGACCCCGG CCGCCTCCGC AAGGAGAGGC	120
GGGGGTCCGAG TCGCCCTGTC CAAAGGCCTC AATCTAAACA ATCTTGATTC CTGTTGCCGG	180
CTGGCGGGAC CCTGAATGGC AGGAAATCTC ACCACATCTC TTCTCCTATC TCCAAGGACC	240
ATG ACC TTG GGG AGC TTC AGA AGG GGC CTT TTG ATG CTG TCG GTG GCC	288
Met Thr Leu Gly Ser Phe Arg Arg Gly Leu Leu Met Leu Ser Val Ala	
1 5 10 15	

TTG GGC CTA ACC CAG GGG AGA CTT GCG AAG CCT TCC AAG CTG GTG AAC	336
Leu Gly Leu Thr 20 Gln Gly Arg Leu Ala Lys Pro Ser Lys Leu Val Asn	25 30
TGC ACT TGT GAG AGC CCA CAC TGC AAG AGA CCA TTC TGC CAG GGG TCA	384
Cys Thr Cys 35 Glu Ser Pro His Cys Lys Arg Pro Phe Cys Gln Gly Ser	40 45
TGG TGC ACA GTG GTG CTG GTT CGA GAG CAG GGC AGG CAC CCC CAG GTC	432
Trp Cys Thr Val Val Leu Val Arg Glu Gln Gly Arg His Pro Gln Val	50 55 60
TAT CCG GGC TGT GGG AGC CTG AAC CAG GAG CTC TGC TTG GGA CGT CCC	480
Tyr Arg Gly Cys Gly Ser Leu Asn Gln Glu Leu Cys Leu Gly Arg Pro	65 70 75 80
ACG GAG TTT CTG AAC CAT CAC TGC TGC TAT AGA TCC TTC TGC AAC CAC	528
Thr Glu Phe Leu Asn His His Cys Cys Tyr Arg Ser Phe Cys Asn His	85 90 95
AAC GTG TCT CTG ATG CTG GAG GCC ACC CAA ACT CCT TCG GAG GAG CCA	576
Asn Val Ser Leu Met Leu Glu Ala Thr Gln Thr Pro Ser Glu Glu Pro	100 105 110
GAA GTT GAT GCC CAT CTG CCT CTG ATC CTG GGT CCT GTG CTG GCC TTG	624
Glu Val Asp Ala His Leu Pro Leu Ile Leu Gly Pro Val Leu Ala Leu	115 120 125
CCG GTC CTG GTG GCC CTG GGT GCT CTG GGC TTG TGG CGT GTC CCG CCG	672
Pro Val Leu Val Ala Leu Gly Ala Leu Gly Leu Trp Arg Val Arg Arg	130 135 140
AGG CAG GAG AAG CAG CCG GAT TTG CAC AGT GAC CTG GGC GAG TCC AGT	720
Arg Gln Glu Lys Gln Arg Asp Leu His Ser Asp Leu Gly Glu Ser Ser	145 150 155 160
CTC ATC CTG AAG GCA TCT GAA CAG GCA GAC AGC ATG TTG GGG GAC TTC	768
Leu Ile Leu Lys Ala Ser Glu Gln Ala Asp Ser Met Leu Gly Asp Phe	165 170 175
CTG GAC AGC GAC TGT ACC ACG GGC AGC GGC TCG GGG CTC CCC TTC TTG	816
Leu Asp Ser Asp Cys Thr Thr Gly Ser Gly Ser Gly Leu Pro Phe Leu	180 185 190
GTG CAG AGG ACG GTA GCT CCG CAG GTT GCG CTG GTA GAG TGT GTG GGA	864
Val Gln Arg Thr Val Ala Arg Gln Val Ala Leu Val Glu Cys Val Gly	195 200 205
AAG GGC CGA TAT GGC GAG GTG TGG CCG GGT TCG TGG CAT GGC GAA AGC	912
Lys Gly Arg Tyr Gly Glu Val Trp Arg Gly Ser Trp His Gly Glu Ser	210 215 220
GTG GCG GTC AAG ATT TTC TCC TCA CGA GAT GAG CAG TCC TGG TTC CCG	960
Val Ala Val Lys Ile Phe Ser Ser Arg Asp Glu Gln Ser Trp Phe Arg	225 230 235 240
GAG ACG GAG ATC TAC AAC ACA GTT CTG CTT AGA CAC GAC AAC ATC CTA	1008
Glu Thr Glu Ile Tyr Asn Thr Val Leu Leu Arg His Asp Asn Ile Leu	245 250 255

GGC TTC ATC GCC TCC GAC ATG ACT TCG CCG AAC TCG AGC ACG CAG CTG Gly Phe Ile Ala Ser Asp Met Thr Ser Arg Asn Ser Ser Thr Gln Leu 260 265 270	1056
TGG CTC ATC ACC CAC TAC CAT GAA CAC GGC TCC CTC TAT GAC TTT CTG Trp Leu Ile Thr His Tyr His Glu His Gly Ser Leu Tyr Asp Phe Leu 275 280 285	1104
CAG AGG CAG ACG CTG GAG CCC CAG TTG GCC CTG AGG CTA GCT GTG TCC Gln Arg Gln Thr Leu Glu Pro Gln Leu Ala Leu Arg Leu Ala Val Ser 290 295 300	1152
CCG GCC TGC GGC CTG GCG CAC CTA CAT GTG GAG ATC TTT GGC ACT CAA Pro Ala Cys Gly Leu Ala His Leu His Val Glu Ile Phe Gly Thr Gln 305 310 315 320	1200
GGC AAA CCA GCC ATT GCC CAT CGT GAC CTC AAG AGT CGC AAT GTG CTG Gly Lys Pro Ala Ile Ala His Arg Asp Leu Lys Ser Arg Asn Val Leu 325 330 335	1248
GTC AAG AGT AAC TTG CAG TGT TGC ATT GCA GAC CTG GGA CTG GCT GTG Val Lys Ser Asn Leu Gln Cys Cys Ile Ala Asp Leu Gly Leu Ala Val 340 345 350	1296
ATG CAC TCA CAA AGC AAC GAG TAC CTG GAT ATC GGC AAC ACA CCC CGA Met His Ser Gln Ser Asn Glu Tyr Leu Asp Ile Gly Asn Thr Pro Arg 355 360 365	1344
GTG GGT ACC AAA AGA TAC ATG GCA CCC GAG GTG CTG GAT GAG CAC ATC Val Gly Thr Lys Arg Tyr Met Ala Pro Glu Val Leu Asp Glu His Ile 370 375 380	1392
CGC ACA GAC TGC TTT GAG TCG TAC AAG TGG ACA GAC ATC TGG GCC TTT Arg Thr Asp Cys Phe Glu Ser Tyr Lys Trp Thr Asp Ile Trp Ala Phe 385 390 395 400	1440
GGC CTA GTG CTA TGG GAG ATC GCC CGG CCG ACC ATC ATC AAT GGC ATT Gly Leu Val Leu Trp Glu Ile Ala Arg Arg Thr Ile Ile Asn Gly Ile 405 410 415	1488
GTG GAG GAT TAC AGG CCA CCT TTC TAT GAC ATG GTA CCC AAT GAC CCC Val Glu Asp Tyr Arg Pro Pro Phe Tyr Asp Met Val Pro Asn Asp Pro 420 425 430	1536
AGT TTT GAG GAC ATG AAA AAG GTG GTG TGC GTT GAC CAG CAG ACA CCC Ser Phe Glu Asp Met Lys Lys Val Val Cys Val Asp Gln Gln Thr Pro 435 440 445	1584
ACC ATC CCT AAC CGG CTG GCT GCA GAT CCG GTC CTC TCC GGG CTG GCC Thr Ile Pro Asn Arg Leu Ala Ala Asp Pro Val Leu Ser Gly Leu Ala 450 455 460	1632
CAG ATG ATG AGA GAG TGC TGG TAC CCC AAC CCC TCT GCT CGC CTC ACC Gln Met Met Arg Glu Cys Trp Tyr Pro Asn Pro Ser Ala Arg Leu Thr 465 470 475 480	1680
GCA CTG CGC ATA AAG AAG ACA TTG CAG AAG CTC AGT CAC AAT CCA GAG Ala Leu Arg Ile Lys Lys Thr Leu Gln Lys Leu Ser His Asn Pro Glu 485 490 495	1728

AAG CCC AAA GTG ATT CAC TAGCCCAGGG CCACCAGGCT TCCTCTGCCT 1776
 Lys Pro Lys Val Ile His
 500

AAAGTGTGTG CTGGGGAAGA AGACATAGCC TGTCTGGGTA GAGGGAGTGA AGAGAGTGTG 1836
 CACGCTGCCC TGTGTGTGCC TGCTCAGCTT GCTCCCAGCC CATCCAGCCA AAAATACAGC 1896
 TGAGCTGAAA TTCAAAAAA AAAAAA 1922

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 502 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Met Thr Leu Gly Ser Phe Arg Arg Gly Leu Leu Met Leu Ser Val Ala
 1 5 10 15

Leu Gly Leu Thr Gln Gly Arg Leu Ala Lys Pro Ser Lys Leu Val Asn
 20 25 30

Cys Thr Cys Glu Ser Pro His Cys Lys Arg Pro Phe Cys Gln Gly Ser
 35 40 45

Trp Cys Thr Val Val Leu Val Arg Glu Gln Gly Arg His Pro Gln Val
 50 55 60

Tyr Arg Gly Cys Gly Ser Leu Asn Gln Glu Leu Cys Leu Gly Arg Pro
 65 70 75 80

Thr Glu Phe Leu Asn His His Cys Cys Tyr Arg Ser Phe Cys Asn His
 85 90 95

Asn Val Ser Leu Met Leu Glu Ala Thr Gln Thr Pro Ser Glu Glu Pro
 100 105 110

Glu Val Asp Ala His Leu Pro Leu Ile Leu Gly Pro Val Leu Ala Leu
 115 120 125

Pro Val Leu Val Ala Leu Gly Ala Leu Gly Leu Trp Arg Val Arg Arg
 130 135 140

Arg Gln Glu Lys Gln Arg Asp Leu His Ser Asp Leu Gly Glu Ser Ser
 145 150 155 160

Leu Ile Leu Lys Ala Ser Glu Gln Ala Asp Ser Met Leu Gly Asp Phe
 165 170 175

Leu Asp Ser Asp Cys Thr Thr Gly Ser Gly Ser Gly Leu Pro Phe Leu
 180 185 190

Val Gln Arg Thr Val Ala Arg Gln Val Ala Leu Val Glu Cys Val Gly
 195 200 205
 Lys Gly Arg Tyr Gly Glu Val Trp Arg Gly Ser Trp His Gly Glu Ser
 210 215 220
 Val Ala Val Lys Ile Phe Ser Ser Arg Asp Glu Gln Ser Trp Phe Arg
 225 230 235 240
 Glu Thr Glu Ile Tyr Acn Thr Val Leu Leu Arg His Asp Asn Ile Leu
 245 250 255
 Gly Phe Ile Ala Ser Asp Met Thr Ser Arg Asn Ser Ser Thr Gln Leu
 260 265 270
 Trp Leu Ile Thr His Tyr His Glu His Gly Ser Leu Tyr Asp Phe Leu
 275 280 285
 Gln Arg Gln Thr Leu Glu Pro Gln Leu Ala Leu Arg Leu Ala Val Ser
 290 295 300
 Pro Ala Cys Gly Leu Ala His Leu His Val Glu Ile Phe Gly Thr Gln
 305 310 315 320
 Gly Lys Pro Ala Ile Ala His Arg Asp Leu Lys Ser Arg Asn Val Leu
 325 330 335
 Val Lys Ser Asn Leu Gln Cys Cys Ile Ala Asp Leu Gly Leu Ala Val
 340 345 350
 Met His Ser Gln Ser Asn Glu Tyr Leu Asp Ile Gly Asn Thr Pro Arg
 355 360 365
 Val Gly Thr Lys Arg Tyr Met Ala Pro Glu Val Leu Asp Glu His Ile
 370 375 380
 Arg Thr Asp Cys Phe Glu Ser Tyr Lys Trp Thr Asp Ile Trp Ala Phe
 385 390 395 400
 Gly Leu Val Leu Trp Glu Ile Ala Arg Arg Thr Ile Ile Asn Gly Ile
 405 410 415
 Val Glu Asp Tyr Arg Pro Pro Phe Tyr Asp Met Val Pro Asn Asp Pro
 420 425 430
 Ser Phe Glu Asp Met Lys Lys Val Val Cys Val Asp Gln Gln Thr Pro
 435 440 445
 Thr Ile Pro Asn Arg Leu Ala Ala Asp Pro Val Leu Ser Gly Leu Ala
 450 455 460
 Gln Met Met Arg Glu Cys Trp Tyr Pro Asn Pro Ser Ala Arg Leu Thr
 465 470 475 480
 Ala Leu Arg Ile Lys Lys Thr Leu Gln Lys Leu Ser His Asn Pro Glu
 485 490 495
 Lys Pro Lys Val Ile His
 500

(2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2070 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cdna

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Mouse

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 217..1812

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

ATT	CAT	GAGA	TGGA	AGCATA	GGT	CAA	AGCT	GTTC	GGAGAA	ATT	GGAA	ACTA	CAG	TTTTATC	60	
TAG	CCACATC	TCT	GAGA	AATT	CTGA	AAG	AAG	CAG	CAGGTGA	AAG	T	CATTGC	CAAG	TGATTT	120	
TG	TTCTGTAA	GGA	AGCCTCC	CTC	ATTCACT	TAC	ACC	CAGTG	AGAC	CAGCAGG	ACC	AGTCATT	180			
CAA	AGGGCCG	TGT	ACAGGAC	GCG	TGGCAAT	CAG	ACA	ATG	ACT	CAG	CTA	TAC	ACT	234		
									Met	Thr	Gln	Leu	Tyr	Thr		
									1				5			
TAC	ATC	AGA	TTA	CTG	GGA	GCC	TGT	CTG	TTC	ATC	ATT	TCT	CAT	GTT	CAA	282
Tyr	Ile	Arg	Leu	Leu	Gly	Ala	Cys	Leu	Phe	Ile	Ile	Ser	His	Val	Gln	
			10					15					20			
GGG	CAG	AAT	CTA	GAT	AGT	ATG	CTC	CAT	GGC	ACT	GGT	ATG	AAA	TCA	GAC	330
Gly	Gln	Asn	Leu	Asp	Ser	Met	Leu	His	Gly	Thr	Gly	Met	Lys	Ser	Asp	
		25					30					35				
TTG	GAC	CAG	AAG	AAG	CCA	GAA	AAT	GGA	GTG	ACT	TTA	GCA	CCA	GAG	GAT	378
Leu	Asp	Gln	Lys	Lys	Pro	Glu	Asn	Gly	Val	Thr	Leu	Ala	Pro	Glu	Asp	
	40					45					50					
ACC	TTG	CCT	TTC	TTA	AAG	TGC	TAT	TGC	TCA	GGA	CAC	TGC	CCA	GAT	GAT	426
Thr	Leu	Pro	Phe	Leu	Lys	Cys	Tyr	Cys	Ser	Gly	His	Cys	Pro	Asp	Asp	
	55				60					65					70	
GCT	ATT	AAT	AAC	ACA	TGC	ATA	ACT	AAT	GGC	CAT	TGC	TTT	GCC	ATT	ATA	474
Ala	Ile	Asn	Asn	Thr	Cys	Ile	Thr	Asn	Gly	His	Cys	Phe	Ala	Ile	Ile	
				75					80					85		
GAA	GAA	GAT	GAT	CAG	GGA	GAA	ACC	ACA	TTA	ACT	TCT	GGG	TGT	ATG	AAG	522
Glu	Glu	Asp	Asp	Gln	Gly	Glu	Thr	Thr	Leu	Thr	Ser	Gly	Cys	Met	Lys	
				90				95					100			

TAT	GAA	GGC	TCT	GAT	TTT	CAA	TGC	AAG	GAT	TCA	CCG	AAA	GCC	CAG	CTA	570
Tyr	Glu	Gly	Ser	Asp	Phe	Gln	Cys	Lys	Asp	Ser	Pro	Lys	Ala	Gln	Leu	
		105					110					115				
CGC	AGG	ACA	ATA	GAA	TGT	TGT	CGG	ACC	AAT	TTG	TGC	AAC	CAG	TAT	TTG	618
Arg	Arg	Thr	Ile	Glu	Cys	Cys	Arg	Thr	Asn	Leu	Cys	Asn	Gln	Tyr	Leu	
		120				125					130					
CAG	CCT	ACA	CTG	CCC	CCT	GTT	GTT	ATA	GGT	CCG	TTC	TTT	GAT	GGC	AGC	666
Gln	Pro	Thr	Leu	Pro	Pro	Val	Val	Ile	Gly	Pro	Phe	Phe	Asp	Gly	Ser	
					140					145					150	
ATC	CGA	TGG	CTG	GTT	GTG	CTC	ATT	TCC	ATG	GCT	GTC	TGT	ATA	GTT	GCT	714
Ile	Arg	Trp	Leu	Val	Val	Leu	Ile	Ser	Met	Ala	Val	Cys	Ile	Val	Ala	
				155					160					165		
ATG	ATC	ATC	TTC	TCC	AGC	TGC	TTT	TGC	TAT	AAG	CAT	TAT	TGT	AAG	AGT	762
Met	Ile	Ile	Phe	Ser	Ser	Cys	Phe	Cys	Tyr	Lys	His	Tyr	Cys	Lys	Ser	
			170					175					180			
ATC	TCA	AGC	AGG	GGT	CGT	TAC	AAC	CGT	GAT	TTG	GAA	CAG	GAT	GAA	GCA	810
Ile	Ser	Ser	Arg	Gly	Arg	Tyr	Asn	Arg	Asp	Leu	Glu	Gln	Asp	Glu	Ala	
		185					190					195				
TTT	ATT	CCA	GTA	GGA	GAA	TCA	TTG	AAA	GAC	CTG	ATT	GAC	CAG	TCC	CAA	858
Phe	Ile	Pro	Val	Gly	Glu	Ser	Leu	Lys	Asp	Leu	Ile	Asp	Gln	Ser	Gln	
		200				205					210					
AGC	TCT	GGG	AGT	GGA	TCT	GGA	TTG	CCT	TTA	TTG	GTT	CAG	CGA	ACT	ATT	906
Ser	Ser	Gly	Ser	Gly	Ser	Gly	Leu	Pro	Leu	Leu	Val	Gln	Arg	Thr	Ile	
		215			220					225					230	
GCC	AAA	CAG	ATT	CAG	ATG	GTT	CGG	CAG	GTT	GGT	AAA	GGC	CGC	TAT	GGA	954
Ala	Lys	Gln	Ile	Gln	Met	Val	Arg	Gln	Val	Gly	Lys	Gly	Arg	Tyr	Gly	
			235					240						245		
GAA	GTA	TGG	ATG	GGT	AAA	TGG	CGT	GGT	GAA	AAA	GTG	GCT	GTC	AAA	GTG	1002
Glu	Val	Trp	Met	Gly	Lys	Trp	Arg	Gly	Glu	Lys	Val	Ala	Val	Lys	Val	
			250					255					260			
TTT	TTT	ACC	ACT	GAA	GAA	GCT	AGC	TGG	TTT	AGA	GAA	ACA	GAA	ATC	TAC	1050
Phe	Phe	Thr	Thr	Glu	Glu	Ala	Ser	Trp	Phe	Arg	Glu	Thr	Glu	Ile	Tyr	
		265					270					275				
CAG	ACG	GTG	TTA	ATG	CGT	CAT	GAA	AAT	ATA	CTT	GGT	TTT	ATA	GCT	GCA	1098
Gln	Thr	Val	Leu	Met	Arg	His	Glu	Asn	Ile	Leu	Gly	Phe	Ile	Ala	Ala	
		280				285					290					
GAC	ATT	AAA	GGC	ACT	GGT	TCC	TGG	ACT	CAG	CTG	TAT	TTG	ATT	ACT	GAT	1146
Asp	Ile	Lys	Gly	Thr	Gly	Ser	Trp	Thr	Gln	Leu	Tyr	Leu	Ile	Thr	Asp	
			295			300				305					310	
TAC	CAT	GAA	AAT	GGA	TCT	CTC	TAT	GAC	TTC	CTG	AAA	TGT	GCC	ACA	CTA	1194
Tyr	His	Glu	Asn	Gly	Ser	Leu	Tyr	Asp	Phe	Leu	Lys	Cys	Ala	Thr	Leu	
			315						320					325		
GAC	ACC	AGA	GCC	CTA	CTC	AAG	TTA	GCT	TAT	TCT	GCT	GCT	TGT	GGT	CTG	1242
Asp	Thr	Arg	Ala	Leu	Leu	Lys	Leu	Ala	Tyr	Ser	Ala	Ala	Cys	Gly	Leu	
			330					335					340			

TGC CAC CTC CAC ACA GAA ATT TAT GGT ACC CAA GGG AAG CCT GCA ATT	1290
Cys His Leu His Thr Glu Ile Tyr Gly Thr Gln Gly Lys Pro Ala Ile	
345 350 355	
GCT CAT CGA GAC CTG AAG AGC AAA AAC ATC CTT ATT AAG AAA AAT GGA	1338
Ala His Arg Asp Leu Lys Ser Lys Asn Ile Leu Ile Lys Lys Asn Gly	
360 365 370	
AGT TGC TGT ATT GCT GAC CTG GGC CTA GCT GTT AAA TTC AAC AGT GAT	1386
Ser Cys Cys Ile Ala Asp Leu Gly Leu Ala Val Lys Phe Asn Ser Asp	
375 380 385 390	
ACA AAT GAA GTT GAC ATA CCC TTG AAT ACC AGG GTG GGC ACC AAG CGG	1434
Thr Asn Glu Val Asp Ile Pro Leu Asn Thr Arg Val Gly Thr Lys Arg	
395 400 405	
TAC ATG GCT CCA GAA GTG CTG GAT GAA AGC CTG AAT AAA AAC CAT TTC	1482
Tyr Met Ala Pro Glu Val Leu Asp Glu Ser Leu Asn Lys Asn His Phe	
410 415 420	
CAG CCC TAC ATC ATG GCT GAC ATC TAT AGC TTT GGT TTG ATC ATT TGG	1530
Gln Pro Tyr Ile Met Ala Asp Ile Tyr Ser Phe Gly Leu Ile Ile Trp	
425 430 435	
GAA ATG GCT CGT CGT TGT ATT ACA GGA GGA ATC GTG GAG GAA TAT CAA	1578
Glu Met Ala Arg Arg Cys Ile Thr Gly Gly Ile Val Glu Glu Tyr Gln	
440 445 450	
TTA CCA TAT TAC AAC ATG GTG CCC AGT GAC CCA TCC TAT GAG GAC ATG	1626
Leu Pro Tyr Tyr Asn Met Val Pro Ser Asp Pro Ser Tyr Glu Asp Met	
455 460 465 470	
CGT GAG GTT GTG TGT GTG AAA CGC TTG CGG CCA ATC GTG TCT AAC CGC	1674
Arg Glu Val Val Cys Val Lys Arg Leu Arg Pro Ile Val Ser Asn Arg	
475 480 485	
TGG AAC AGC GAT GAA TGT CTT CGA GCA GTT TTG AAG CTA ATG TCA GAA	1722
Trp Asn Ser Asp Glu Cys Leu Arg Ala Val Leu Lys Leu Met Ser Glu	
490 495 500	
TGT TGG GCC CAT AAT CCA GCC TCC AGA CTC ACA GCT TTG AGA ATC AAG	1770
Cys Trp Ala His Asn Pro Ala Ser Arg Leu Thr Ala Leu Arg Ile Lys	
505 510 515	
AAG ACA CTT GCA AAA ATG GTT GAA TCC CAG GAT GTA AAG ATT	1812
Lys Thr Leu Ala Lys Met Val Glu Ser Gln Asp Val Lys Ile	
520 525 530	
TGACAATTAA ACAATTTTGA GGGAGAATTT AGACTGCAAG AACTTCTTCA CCCAAGGAAT	1872
GGGTGGGATT AGCATGGAAT AGGATGTTGA CTTGGTTTCC AGACTCCTTC CTCTACATCT	1932
TCACAGGCTG CTAACAGTAA ACCTTACCGT ACTCTACAGA ATACAAGATT GGAACCTTGGG	1992
ACTTCAAACA TGTCATTCTT TATATATGAC AGCTTTGTTT TAATGTGGGG TTTTTTTGTT	2052
TGCTTTTTTT GTTTTGTT	2070

(2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 532 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

Met Thr Gln Leu Tyr Thr Tyr Ile Arg Leu Leu Gly Ala Cys Leu Phe
 1 5 10 15
 Ile Ile Ser His Val Gln Gly Gln Asn Leu Asp Ser Met Leu His Gly
 20 25 30
 Thr Gly Met Lys Ser Asp Leu Asp Gln Lys Lys Pro Glu Asn Gly Val
 35 40 45
 Thr Leu Ala Pro Glu Asp Thr Leu Pro Phe Leu Lys Cys Tyr Cys Ser
 50 55 60
 Gly His Cys Pro Asp Asp Ala Ile Asn Asn Thr Cys Ile Thr Asn Gly
 65 70 75 80
 His Cys Phe Ala Ile Ile Glu Glu Asp Asp Gln Gly Glu Thr Thr Leu
 85 90 95
 Thr Ser Gly Cys Met Lys Tyr Glu Gly Ser Asp Phe Gln Cys Lys Asp
 100 105 110
 Ser Pro Lys Ala Gln Leu Arg Arg Thr Ile Glu Cys Cys Arg Thr Asn
 115 120 125
 Leu Cys Asn Gln Tyr Leu Gln Pro Thr Leu Pro Pro Val Val Ile Gly
 130 135 140
 Pro Phe Phe Asp Gly Ser Ile Arg Trp Leu Val Val Leu Ile Ser Met
 145 150 155 160
 Ala Val Cys Ile Val Ala Met Ile Ile Phe Ser Ser Cys Phe Cys Tyr
 165 170 175
 Lys His Tyr Cys Lys Ser Ile Ser Ser Arg Gly Arg Tyr Asn Arg Asp
 180 185 190
 Leu Glu Gln Asp Glu Ala Phe Ile Pro Val Gly Glu Ser Leu Lys Asp
 195 200 205
 Leu Ile Asp Gln Ser Gln Ser Ser Gly Ser Gly Ser Gly Leu Pro Leu
 210 215 220
 Leu Val Gln Arg Thr Ile Ala Lys Gln Ile Gln Met Val Arg Gln Val
 225 230 235 240
 Gly Lys Gly Arg Tyr Gly Glu Val Trp Met Gly Lys Trp Arg Gly Glu
 245 250 255

SUBSTITUTE SHEET.

Lys Val Ala Val Lys Val Phe Phe Thr Thr Glu Glu Ala Ser Trp Phe
 260 265 270
 Arg Glu Thr Glu Ile Tyr Gln Thr Val Leu Met Arg His Glu Asn Ile
 275 280 285
 Leu Gly Phe Ile Ala Ala Asp Ile Lys Gly Thr Gly Ser Trp Thr Gln
 290 295 300
 Leu Tyr Leu Ile Thr Asp Tyr His Glu Asn Gly Ser Leu Tyr Asp Phe
 305 310 315 320
 Leu Lys Cys Ala Thr Leu Asp Thr Arg Ala Leu Leu Lys Leu Ala Tyr
 325 330 335
 Ser Ala Ala Cys Gly Leu Cys His Leu His Thr Glu Ile Tyr Gly Thr
 340 345 350
 Gln Gly Lys Pro Ala Ile Ala His Arg Asp Leu Lys Ser Lys Asn Ile
 355 360 365
 Leu Ile Lys Lys Asn Gly Ser Cys Cys Ile Ala Asp Leu Gly Leu Ala
 370 375 380
 Val Lys Phe Asn Ser Asp Thr Asn Glu Val Asp Ile Pro Leu Asn Thr
 385 390 395 400
 Arg Val Gly Thr Lys Arg Tyr Met Ala Pro Glu Val Leu Asp Glu Ser
 405 410 415
 Leu Asn Lys Asn His Phe Gln Pro Tyr Ile Met Ala Asp Ile Tyr Ser
 420 425 430
 Phe Gly Leu Ile Ile Trp Glu Met Ala Arg Arg Cys Ile Thr Gly Gly
 435 440 445
 Ile Val Glu Glu Tyr Gln Leu Pro Tyr Tyr Asn Met Val Pro Ser Asp
 450 455 460
 Pro Ser Tyr Glu Asp Met Arg Glu Val Val Cys Val Lys Arg Leu Arg
 465 470 475 480
 Pro Ile Val Ser Asn Arg Trp Asn Ser Asp Glu Cys Leu Arg Ala Val
 485 490 495
 Leu Lys Leu Met Ser Glu Cys Trp Ala His Asn Pro Ala Ser Arg Leu
 500 505 510
 Thr Ala Leu Arg Ile Lys Lys Thr Leu Ala Lys Met Val Glu Ser Gln
 515 520 525
 Asp Val Lys Ile
 530

(2) INFORMATION FOR SEQ ID NO: 15:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2160 base pairs

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: Mouse

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 10..1524

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CGCGGTTAC ATG GCG GAG TCG GCC GGA GCC TCC TCC TTC TTC CCC CTT	48
Met Ala Glu Ser Ala Gly Ala Ser Ser Phe Phe Pro Leu	
1 5 10	
GTT GTC CTC CTG CTC GCC GGC AGC GGC GGG TCC GGG CCC CGG GGG ATC	96
Val Val Leu Leu Leu Ala Gly Ser Gly Gly Ser Gly Pro Arg Gly Ile	
15 20 25	
CAG GCT CTG CTG TGT GCG TGC ACC AGC TGC CTA CAG ACC AAC TAC ACC	144
Gln Ala Leu Leu Cys Ala Cys Thr Ser Cys Leu Gln Thr Asn Tyr Thr	
30 35 40 45	
TGT GAG ACA GAT GGG GCT TGC ATG GTC TCC ATC TTT AAC CTG GAT GGC	192
Cys Glu Thr Asp Gly Ala Cys Met Val Ser Ile Phe Asn Leu Asp Gly	
50 55 60	
GTG GAG CAC CAT GTA CGT ACC TGC ATC CCC AAG GTG GAG CTG GTT CCT	240
Val Glu His His Val Arg Thr Cys Ile Pro Lys Val Glu Leu Val Pro	
65 70 75	
GCT GGA AAG CCC TTC TAC TGC CTG AGT TCA GAG GAT CTG CGC AAC ACA	288
Ala Gly Lys Pro Phe Tyr Cys Leu Ser Ser Glu Asp Leu Arg Asn Thr	
80 85 90	
CAC TGC TGC TAT ATT GAC TTC TGC AAC AAG ATT GAC CTC AGG GTC CCC	336
His Cys Cys Tyr Ile Asp Phe Cys Asn Lys Ile Asp Leu Arg Val Pro	
95 100 105	
AGC GGA CAC CTC AAG GAG CCT GCG CAC CCC TCC ATG TGG GGC CCT GTG	384
Ser Gly His Leu Lys Glu Pro Ala His Pro Ser Met Trp Gly Pro Val	
110 115 120 125	
GAG CTG GTC GGC ATC ATC GCC GGC CCC GTC TTC CTC CTC TTC CTT ATC	432
Glu Leu Val Gly Ile Ile Ala Gly Pro Val Phe Leu Leu Phe Leu Ile	
130 135 140	

ATT	ATC	ATC	GTC	TTC	CTG	GTC	ATC	AAC	TAT	CAC	CAG	CCT	GTC	TAC	CAT	480
Ile	Ile	Ile	Val	Phe	Leu	Val	Ile	Asn	Tyr	His	Gln	Arg	Val	Tyr	His	
			145					150					155			
AAC	CGC	CAG	AGG	TTG	GAC	ATG	GAG	GAC	CCC	TCT	TGC	GAG	ATG	TGT	CTC	528
Asn	Arg	Gln	Arg	Leu	Asp	Met	Glu	Asp	Pro	Ser	Cys	Glu	Met	Cys	Leu	
		160					165					170				
TCC	AAA	GAC	AAG	ACG	CTC	CAG	GAT	CTC	GTC	TAC	GAC	CTC	TCC	ACG	TCA	576
Ser	Lys	Asp	Lys	Thr	Leu	Gln	Asp	Leu	Val	Tyr	Asp	Leu	Ser	Thr	Ser	
	175					180					185					
GGG	TCT	GGC	TCA	GGG	TTA	CCC	CTT	TTT	GTC	CAG	CGC	ACA	GTG	GCC	CGA	624
Gly	Ser	Gly	Ser	Gly	Leu	Pro	Leu	Phe	Val	Gln	Arg	Thr	Val	Ala	Arg	
190				195				200					205			
ACC	ATT	GTT	TTA	CAA	GAG	ATT	ATC	GGC	AAG	GGC	CGG	TTC	GGG	GAA	GTA	672
Thr	Ile	Val	Leu	Gln	Glu	Ile	Ile	Gly	Lys	Gly	Arg	Phe	Gly	Glu	Val	
			210					215					220			
TGG	CGT	GGT	CGC	TGG	AGG	GGT	GGT	GAC	GTG	GCT	GTG	AAA	ATC	TTC	TCT	720
Trp	Arg	Gly	Arg	Trp	Arg	Gly	Gly	Asp	Val	Ala	Val	Lys	Ile	Phe	Ser	
			225					230					235			
TCT	CGT	GAA	GAA	CGG	TCT	TGG	TTC	CGT	GAA	GCA	GAG	ATC	TAC	CAG	ACC	768
Ser	Arg	Glu	Glu	Arg	Ser	Trp	Phe	Arg	Glu	Ala	Glu	Ile	Tyr	Gln	Thr	
		240					245					250				
GTC	ATG	CTG	CGC	CAT	GAA	AAC	ATC	CTT	GGC	TTT	ATT	GCT	GCT	GAC	AAT	816
Val	Met	Leu	Arg	His	Glu	Asn	Ile	Leu	Gly	Phe	Ile	Ala	Ala	Asp	Asn	
	255					260					265					
AAA	GAT	AAT	GGC	ACC	TGG	ACC	CAG	CTG	TGG	CTT	GTC	TCT	GAC	TAT	CAC	864
Lys	Asp	Asn	Gly	Thr	Trp	Thr	Gln	Leu	Trp	Leu	Val	Ser	Asp	Tyr	His	
270				275				280						285		
GAG	CAT	GGC	TCA	CTG	TTT	GAT	TAT	CTG	AAC	CGC	TAC	ACA	GTG	ACC	ATT	912
Glu	His	Gly	Ser	Leu	Phe	Asp	Tyr	Leu	Asn	Arg	Tyr	Thr	Val	Thr	Ile	
			290					295						300		
GAG	GGA	ATG	ATT	AAG	CTA	GCC	TTG	TCT	GCA	GCC	AGT	GGT	TTG	GCA	CAC	960
Glu	Gly	Met	Ile	Lys	Leu	Ala	Leu	Ser	Ala	Ala	Ser	Gly	Leu	Ala	His	
		305					310						315			
CTG	CAT	ATG	GAG	ATT	GTG	GGC	ACT	CAA	GGG	AAG	CCG	GGA	ATT	GCT	CAT	1008
Leu	His	Met	Glu	Ile	Val	Gly	Thr	Gln	Gly	Lys	Pro	Gly	Ile	Ala	His	
		320				325						330				
CGA	GAC	TTG	AAG	TCA	AAG	AAC	ATC	CTG	GTG	AAA	AAA	AAT	GGC	ATG	TGT	1056
Arg	Asp	Leu	Lys	Ser	Lys	Asn	Ile	Leu	Val	Lys	Lys	Asn	Gly	Met	Cys	
	335					340					345					
GCC	ATT	GCA	GAC	CTG	GGC	CTG	GCT	GTC	CGT	CAT	GAT	GCG	GTC	ACT	GAC	1104
Ala	Ile	Ala	Asp	Leu	Gly	Leu	Ala	Val	Arg	His	Asp	Ala	Val	Thr	Asp	
350				355				360						365		
ACC	ATA	GAC	ATT	GCT	CCA	AAT	CAG	AGG	GTG	GGG	ACC	AAA	CGA	TAC	ATG	1152
Thr	Ile	Asp	Ile	Ala	Pro	Asn	Gln	Arg	Val	Gly	Thr	Lys	Arg	Tyr	Met	
				370				375						380		

GCT CCT GAA GTC CTT GAC GAG ACA ATC AAC ATG AAG CAC TTT GAC TCC	1200
Ala Pro Glu Val Leu Asp Glu Thr Ile Asn Met Lys His Phe Asp Ser	
385 390 395	
TTC AAA TGT GCC GAC ATC TAT GCC CTC GGG CTT GTC TAC TGG GAG ATT	1248
Phe Lys Cys Ala Asp Ile Tyr Ala Leu Gly Leu Val Tyr Trp Glu Ile	
400 405 410	
GCA CGA AGA TGC AAT TCT GGA GGA GTC CAT GAA GAC TAT CAA CTG CCG	1296
Ala Arg Arg Cys Asn Ser Gly Gly Val His Glu Asp Tyr Gln Leu Pro	
415 420 425	
TAT TAC GAC TTA GTG CCC TCC GAC CCT TCC ATT GAG GAG ATG CGA AAG	1344
Tyr Tyr Asp Leu Val Pro Ser Asp Pro Ser Ile Glu Glu Met Arg Lys	
430 435 440 445	
GTT GTA TGT GAC CAG AAG CTA CGG CCC AAT GTC CCC AAC TGG TGG CAG	1392
Val Val Cys Asp Gln Lys Leu Arg Pro Asn Val Pro Asn Trp Trp Gln	
450 455 460	
AGT TAT GAG GCC TTG CGA GTG ATG GGA AAG ATG ATG CGG GAG TGC TGG	1440
Ser Tyr Glu Ala Leu Arg Val Met Gly Lys Met Met Arg Glu Cys Trp	
465 470 475	
TAC GCC AAT GGT GCT GCC CGT CTG ACA GCT CTG CGC ATC AAG AAG ACT	1488
Tyr Ala Asn Gly Ala Ala Arg Leu Thr Ala Leu Arg Ile Lys Lys Thr	
480 485 490	
CTG TCC CAG CTA AGC GTG CAG GAA GAT GTG AAG ATT TAAGCTGTTT	1534
Leu Ser Gln Leu Ser Val Gln Glu Asp Val Lys Ile	
495 500 505	
CTCTGCCTAC ACAAAGAACC TGGGCAGTGA GGATGACTGC AGCCACCGTG CAAGCGTGT	1594
GGAGGCCTAT CCTCTTGTIT CTGCCCCGGCC CTCTGGCAGA GCCCTGGCCT GCAAGAGGGA	1654
CAGAGCCTGG GAGACGCGCG CACTCCCCTT GGGTTTGTAGA CAGACACTTT TTATATTTAC	1714
CTCCTGATGG CATGGAGACC TGAGCAAATC ATGTAGTCAC TCAATGCCAC AACTCAAAC	1774
GCTTCAGTGG GAAGTACAGA GACCCAGTGC ATTGCGTGTG CAGGAGCGTG AGGTGCTGGG	1834
CTCGCCAGGA GCGGCCCCCA TACCTTGTGG TCCACTGGGC TGCAGGTTTT CCTCCAGGGA	1894
CCAGTCAACT GGCATCAAGA TATTGAGAGG AACCGGAAGT TTCTCCCTCC TTCCCGTAGC	1954
AGTCCTGAGC CACACCATCC TTCTCATGGA CATCCGGAGG ACTGCCCCTA GAGACACAAC	2014
CTGCTGCCTG TCTGTCCAGC CAAGTGCCGA TGTGCCGAGG TGTGTCCAC ATTGTGCCTG	2074
GTCTGTGCCA CGCCCGTGTG TGTGTGTGTG TGTGTGAGTG AGTGTGTGTG TGTACACTTA	2134
ACCTGCTTGA GCTTCTGTGC ATGTGT	2160

(2) INFORMATION FOR SEQ ID NO: 16:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 505 amino acids

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

```

Met Ala Glu Ser Ala Gly Ala Ser Ser Phe Phe Pro Leu Val Val Leu
 1          5          10          15
Leu Leu Ala Gly Ser Gly Gly Ser Gly Pro Arg Gly Ile Gln Ala Leu
 20          25          30
Leu Cys Ala Cys Thr Ser Cys Leu Gln Thr Asn Tyr Thr Cys Glu Thr
 35          40          45
Asp Gly Ala Cys Met Val Ser Ile Phe Asn Leu Asp Gly Val Glu His
 50          55          60
His Val Arg Thr Cys Ile Pro Lys Val Glu Leu Val Pro Ala Gly Lys
 65          70          75          80
Pro Phe Tyr Cys Leu Ser Ser Glu Asp Leu Arg Asn Thr His Cys Cys
 85          90          95
Tyr Ile Asp Phe Cys Asn Lys Ile Asp Leu Arg Val Pro Ser Gly His
 100          105          110
Leu Lys Glu Pro Ala His Pro Ser Met Trp Gly Pro Val Glu Leu Val
 115          120          125
Gly Ile Ile Ala Gly Pro Val Phe Leu Leu Phe Leu Ile Ile Ile Ile
 130          135          140
Val Phe Leu Val Ile Asn Tyr His Gln Arg Val Tyr His Asn Arg Gln
 145          150          155          160
Arg Leu Asp Met Glu Asp Pro Ser Cys Glu Met Cys Leu Ser Lys Asp
 165          170          175
Lys Thr Leu Gln Asp Leu Val Tyr Asp Leu Ser Thr Ser Gly Ser Gly
 180          185          190
Ser Gly Leu Pro Leu Phe Val Gln Arg Thr Val Ala Arg Thr Ile Val
 195          200          205
Leu Gln Glu Ile Ile Gly Lys Gly Arg Phe Gly Glu Val Trp Arg Gly
 210          215          220
Arg Trp Arg Gly Gly Asp Val Ala Val Lys Ile Phe Ser Ser Arg Glu
 225          230          235          240
Glu Arg Ser Trp Phe Arg Glu Ala Glu Ile Tyr Gln Thr Val Met Leu
 245          250          255
Arg His Glu Asn Ile Leu Gly Phe Ile Ala Ala Asp Asn Lys Asp Asn
 260          265          270

```

74

Gly Thr Trp Thr Gln Leu Trp Leu Val Ser Asp Tyr His Glu His Gly
 275 280 285
 Ser Leu Phe Asp Tyr Leu Asn Arg Tyr Thr Val Thr Ile Glu Gly Met
 290 295 300
 Ile Lys Leu Ala Leu Ser Ala Ala Ser Gly Leu Ala His Leu His Met
 305 310 315 320
 Glu Ile Val Gly Thr Gln Gly Lys Pro Gly Ile Ala His Arg Asp Leu
 325 330 335
 Lys Ser Lys Asn Ile Leu Val Lys Lys Asn Gly Met Cys Ala Ile Ala
 340 345 350
 Asp Leu Gly Leu Ala Val Arg His Asp Ala Val Thr Asp Thr Ile Asp
 355 360 365
 Ile Ala Pro Asn Gln Arg Val Gly Thr Lys Arg Tyr Met Ala Pro Glu
 370 375 380
 Val Leu Asp Glu Thr Ile Asn Met Lys His Phe Asp Ser Phe Lys Cys
 385 390 395 400
 Ala Asp Ile Tyr Ala Leu Gly Leu Val Tyr Trp Glu Ile Ala Arg Arg
 405 410 415
 Cys Asn Ser Gly Gly Val His Glu Asp Tyr Gln Leu Pro Tyr Tyr Asp
 420 425 430
 Leu Val Pro Ser Asp Pro Ser Ile Glu Glu Met Arg Lys Val Val Cys
 435 440 445
 Asp Gln Lys Leu Arg Pro Asn Val Pro Asn Trp Trp Gln Ser Tyr Glu
 450 455 460
 Ala Leu Arg Val Met Gly Lys Met Met Arg Glu Cys Trp Tyr Ala Asn
 465 470 475 480
 Gly Ala Ala Arg Leu Thr Ala Leu Arg Ile Lys Lys Thr Leu Ser Gln
 485 490 495
 Leu Ser Val Gln Glu Asp Val Lys Ile
 500 505

(2) INFORMATION FOR SEQ ID NO: 17:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1952 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Mouse

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 187..1692

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

AAGCGGCGGC AGAAGTTGCC GCGGTGGTGC TCGTAGTGAG GGCGCGGAGG ACCCGGGACC	60
TGGGAAGCGG CGGCGGGTTA ACTTCGGCTG AATCACAACC ATTTGGCGCT GAGCTATGAC	120
AAAGAGGCAA ACAAAAAGTT AAAGGAGCAA CCCGGCCATA AGTGAAGAGA GAAGTTTATT	180
GATAAC ATG CTC TTA CGA AGC TCT GGA AAA TTA AAT GTG GGC ACC AAG	228
Met Leu Leu Arg Ser Ser Gly Lys Leu Asn Val Gly Thr Lys	
1 5 10	
AAG GAG GAT GGA GAG AGT ACA GCC CCC ACC CCT CGG CCC AAG ATC CTA	276
Lys Glu Asp Gly Glu Ser Thr Ala Pro Thr Pro Arg Pro Lys Ile Leu	
15 20 25 30	
CGT TGT AAA TGC CAC CAC CAC TGT CCG GAA GAC TCA GTC AAC AAT ATC	324
Arg Cys Lys Cys His His His Cys Pro Glu Asp Ser Val Asn Asn Ile	
35 40 45	
TGC AGC ACA GAT GGG TAC TGC TTC ACG ATG ATA GAA GAA GAT GAC TCT	372
Cys Ser Thr Asp Gly Tyr Cys Phe Thr Met Ile Glu Glu Asp Asp Ser	
50 55 60	
GGA ATG CCT GTT GTC ACC TCT GGA TGT CTA GGA CTA GAA GGG TCA GAT	420
Gly Met Pro Val Val Thr Ser Gly Cys Leu Gly Leu Glu Gly Ser Asp	
65 70 75	
TTT CAA TGT CGT GAC ACT CCC ATT CCT CAT CAA AGA AGA TCA ATT GAA	468
Phe Gln Cys Arg Asp Thr Pro Ile Pro His Gln Arg Arg Ser Ile Glu	
80 85 90	
TGC TGC ACA GAA AGG AAT GAG TGT AAT AAA GAC CTC CAC CCC ACT CTG	516
Cys Cys Thr Glu Arg Asn Glu Cys Asn Lys Asp Leu His Pro Thr Leu	
95 100 105 110	
CCT CCT CTC AAG GAC AGA GAT TTT GTT GAT GGG CCC ATA CAC CAC AAG	564
Pro Pro Leu Lys Asp Arg Asp Phe Val Asp Gly Pro Ile His His Lys	
115 120 125	
GCC TTG CTT ATC TCT GTG ACT GTC TGT AGT TTA CTC TTG GTC CTC ATT	612
Ala Leu Leu Ile Ser Val Thr Val Cys Ser Leu Leu Leu Val Leu Ile	
130 135 140	
ATT TTA TTC TGT TAC TTC AGG TAT AAA AGA CAA GAA GCC CGA CCT CCG	660
Ile Leu Phe Cys Tyr Phe Arg Tyr Lys Arg Gln Glu Ala Arg Pro Arg	
145 150 155	

TAC Tyr	AGC Ser	ATT Ile	GGG Gly	CTG Leu	GAG Glu	CAG Gln	GAC Asp	GAG Glu	ACA Thr	TAC Tyr	ATT Ile	CCT Pro	CCT Pro	GGA Gly	GAG Glu	708
160						165					170					
TCC Ser	CTG Leu	AGA Arg	GAC Asp	TTG Leu	ATC Ile	GAG Glu	CAG Gln	TCT Ser	CAG Gln	AGC Ser	TCG Ser	GGA Gly	AGT Ser	GGA Gly	TCA Ser	756
175					180					185					190	
GGC Gly	CTC Leu	CCT Pro	CTG Leu	CTG Leu	GTC Val	CAA Gln	AGG Arg	ACA Thr	ATA Ile	GCT Ala	AAG Lys	CAA Gln	ATT Ile	CAG Gln	ATG Met	804
				195					200					205		
GTG Val	AAG Lys	CAG Gln	ATT Ile	GGA Gly	AAA Lys	GGC Gly	CGC Arg	TAT Tyr	GGC Gly	GAG Glu	GTG Val	TCG Trp	ATG Met	GGA Gly	AAG Lys	852
			210					215					220			
TGG Trp	CGT Arg	GGA Gly	GAA Glu	AAG Lys	GTG Val	GCT Ala	GTG Val	AAA Lys	GTG Val	TTC Phe	TTC Phe	ACC Thr	ACG Thr	GAG Glu	GAA Glu	900
		225					230					235				
GCC Ala	AGC Ser	TGG Trp	TTC Phe	CGA Arg	GAG Glu	ACT Thr	GAG Glu	ATA Ile	TAT Tyr	CAG Gln	ACG Thr	GTC Val	CTG Leu	ATG Met	CGG Arg	948
		240				245					250					
CAT His	GAG Glu	AAT Asn	ATT Ile	CTG Leu	GGG Gly	TTC Phe	ATT Ile	GCT Ala	GCA Ala	GAT Asp	ATC Ile	AAA Lys	GGG Gly	ACT Thr	GGG Gly	996
255					260					265					270	
TCC Ser	TGG Trp	ACT Thr	CAG Gln	TTG Leu	TAC Tyr	CTC Leu	ATC Ile	ACA Thr	GAC Asp	TAT Tyr	CAT His	GAA Glu	AAC Asn	GGC Gly	TCC Ser	1044
				275					280					285		
CTT Leu	TAT Tyr	GAC Asp	TAT Tyr	CTG Leu	AAA Lys	TCC Ser	ACC Thr	ACC Thr	TTA Leu	GAC Asp	GCA Ala	AAG Lys	TCC Ser	ATG Met	CTG Leu	1092
				290				295					300			
AAG Lys	CTA Leu	GCC Ala	TAC Tyr	TCC Ser	TCT Ser	GTG Val	AGC Ser	GGC Gly	CTA Leu	TGC Cys	CAT His	TTA Leu	CAC His	ACG Thr	GAA Glu	1140
		305					310					315				
ATC Ile	TTT Phe	AGC Ser	ACT Thr	CAA Gln	GGC Gly	AAG Lys	CCA Pro	GCA Ala	ATC Ile	GCC Ala	CAT His	CGA Arg	GAC Asp	TTG Leu	AAA Lys	1188
	320					325					330					
AGT Ser	AAA Lys	AAC Asn	ATC Ile	CTG Leu	GTG Val	AAG Lys	AAA Lys	AAT Asn	GGA Gly	ACT Thr	TGC Cys	TGC Cys	ATA Ile	GCA Ala	GAC Asp	1236
					340					345					350	
CTG Leu	GGC Gly	TTG Leu	GCT Ala	GTC Val	AAG Lys	TTC Phe	ATT Ile	AGT Ser	GAC Asp	ACA Thr	AAT Asn	GAG Glu	GTT Val	GAC Asp	ATC Ile	1284
				355					360					365		
CCA Pro	CCC Pro	AAC Asn	ACC Thr	CGG Arg	GTT Val	GGC Gly	ACC Thr	AAG Lys	CGC Arg	TAT Tyr	ATG Met	CCT Pro	CCA Pro	GAA Glu	GTG Val	1332
				370				375					380			
CTG Leu	GAC Asp	GAG Glu	AGC Ser	TTG Leu	AAT Asn	AGA Arg	AAC Asn	CAT His	TTC Phe	CAG Gln	TCC Ser	TAC Tyr	ATT Ile	ATG Met	GCT Ala	1380
		385					390					395				

GAC ATG TAC AGC TTT GGA CTC ATC CTC TGG GAG ATT GCA AGG AGA TGT	1428
Asp Met Tyr Ser Phe Gly Leu Ile Leu Trp Glu Ile Ala Arg Arg Cys	
400 405 410	
GTT TCT GGA GGT ATA GTG GAA GAA TAC CAG CTT CCC TAT CAC GAC CTG	1476
Val Ser Gly Gly Ile Val Glu Glu Tyr Gln Leu Pro Tyr His Asp Leu	
415 420 425 430	
GTG CCC AGT GAC CCT TCT TAT GAG GAC ATG AGA GAA ATT GTG TGC ATG	1524
Val Pro Ser Asp Pro Ser Tyr Glu Asp Met Arg Glu Ile Val Cys Met	
435 440 445	
AAG AAG TTA CGG CCT TCA TTC CCC AAT CGA TGG AGC AGT GAT GAG TGT	1572
Lys Lys Leu Arg Pro Ser Phe Pro Asn Arg Trp Ser Ser Asp Glu Cys	
450 455 460	
CTC AGG CAG ATG GGG AAG CTT ATG ACA GAG TGC TGG GCG CAG AAT CCT	1620
Leu Arg Gln Met Gly Lys Leu Met Thr Glu Cys Trp Ala Gln Asn Pro	
465 470 475	
GCC TCC AGG CTG ACG GCC CTG AGA GTT AAG AAA ACC CTT GCC AAA ATG	1668
Ala Ser Arg Leu Thr Ala Leu Arg Val Lys Lys Thr Leu Ala Lys Met	
480 485 490	
TCA GAG TCC CAG GAC ATT AAA CTC TGACGTCAGA TACTTGTGGA CAGAGCAAGA	1722
Ser Glu Ser Gln Asp Ile Lys Leu	
495 500	
ATTTCACAGA AGCATCGTTA GCCCAAGCCT TGAACGTTAG CCTACTGCCC AGTGAGTTCA	1782
GACTTTCCTG GAAGAGAGCA CCGTGGGCAG ACACAGAGGA ACCCAGAAAC ACGGATTCAT	1842
CATGGCTTTC TGAGGAGCAG AAACGTGTTG GGTAACCTGT TCAAGATATG ATGCATGTTG	1902
CTTCTAAGA AAGCCCTGTA TTTTGAATTA CCATTTTTTT ATAAAAAAA	1952

(2) INFORMATION FOR SEQ ID NO: 18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 502 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Met Leu Leu Arg Ser Ser Gly Lys Leu Asn Val Gly Thr Lys Lys Glu	
1 5 10 15	
Asp Gly Glu Ser Thr Ala Pro Thr Pro Arg Pro Lys Ile Leu Arg Cys	
20 25 30	
Lys Cys His His His Cys Pro Glu Asp Ser Val Asn Asn Ile Cys Ser	
35 40 45	
Thr Asp Gly Tyr Cys Phe Thr Met Ile Glu Glu Asp Asp Ser Gly Met	
50 55 60	

78

Pro Val Val Thr Ser Gly Cys Leu Gly Leu Glu Gly Ser Asp Phe Gln
 65 70 75 80
 Cys Arg Asp Thr Pro Ile Pro His Gln Arg Arg Ser Ile Glu Cys Cys
 85 90 95
 Thr Glu Arg Asn Glu Cys Asn Lys Asp Leu His Pro Thr Leu Pro Pro
 100 105 110
 Leu Lys Asp Arg Asp Phe Val Asp Gly Pro Ile His His Lys Ala Leu
 115 120 125
 Leu Ile Ser Val Thr Val Cys Ser Leu Leu Leu Val Leu Ile Ile Leu
 130 135 140
 Phe Cys Tyr Phe Arg Tyr Lys Arg Gln Glu Ala Arg Pro Arg Tyr Ser
 145 150 155 160
 Ile Gly Leu Glu Gln Asp Glu Thr Tyr Ile Pro Pro Gly Glu Ser Leu
 165 170 175
 Arg Asp Leu Ile Glu Gln Ser Gln Ser Ser Gly Ser Gly Ser Gly Leu
 180 185 190
 Pro Leu Leu Val Gln Arg Thr Ile Ala Lys Gln Ile Gln Met Val Lys
 195 200 205
 Gln Ile Gly Lys Gly Arg Tyr Gly Glu Val Trp Met Gly Lys Trp Arg
 210 215 220
 Gly Glu Lys Val Ala Val Lys Val Phe Phe Thr Thr Glu Glu Ala Ser
 225 230 235 240
 Trp Phe Arg Glu Thr Glu Ile Tyr Gln Thr Val Leu Met Arg His Glu
 245 250 255
 Asn Ile Leu Gly Phe Ile Ala Ala Asp Ile Lys Gly Thr Gly Ser Trp
 260 265 270
 Thr Gln Leu Tyr Leu Ile Thr Asp Tyr His Glu Asn Gly Ser Leu Tyr
 275 280 285
 Asp Tyr Leu Lys Ser Thr Thr Leu Asp Ala Lys Ser Met Leu Lys Leu
 290 295 300
 Ala Tyr Ser Ser Val Ser Gly Leu Cys His Leu His Thr Glu Ile Phe
 305 310 315 320
 Ser Thr Gln Gly Lys Pro Ala Ile Ala His Arg Asp Leu Lys Ser Lys
 325 330 335
 Asn Ile Leu Val Lys Lys Asn Gly Thr Cys Cys Ile Ala Asp Leu Gly
 340 345 350
 Leu Ala Val Lys Phe Ile Ser Asp Thr Asn Glu Val Asp Ile Pro Pro
 355 360 365
 Asn Thr Arg Val Gly Thr Lys Arg Tyr Met Pro Pro Glu Val Leu Asp
 370 375 380

SUBSTITUTE SHEET

79

Glu Ser Leu Asn Arg Asn His Phe Gln Ser Tyr Ile Met Ala Asp Met
 385 390 395 400
 Tyr Ser Phe Gly Leu Ile Leu Trp Glu Ile Ala Arg Arg Cys Val Ser
 405 410 415
 Gly Gly Ile Val Glu Glu Tyr Gln Leu Pro Tyr His Asp Leu Val Pro
 420 425 430
 Ser Asp Pro Ser Tyr Glu Asp Met Arg Glu Ile Val Cys Met Lys Lys
 435 440 445
 Leu Arg Pro Ser Phe Pro Asn Arg Trp Ser Ser Asp Glu Cys Leu Arg
 450 455 460
 Gln Met Gly Lys Leu Met Thr Glu Cys Trp Ala Gln Asn Pro Ala Ser
 465 470 475 480
 Arg Leu Thr Ala Leu Arg Val Lys Lys Thr Leu Ala Lys Met Ser Glu
 485 490 495
 Ser Gln Asp Ile Lys Leu
 500

(2) INFORMATION FOR SEQ ID NO: 19:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 28 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

GCGGATCCTG TTGTGAAGGN AATATGTC

28

(2) INFORMATION FOR SEQ ID NO: 20:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

GCGATCCGTC GCAGTCAAAA TTTT

24

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GCGGATCCGC GATATATTAA AAGCAA

26

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

CGGAATTCG GTGCCATATA

20

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

ATTCAAGGGC ACATCAACTT CATTGTGTC ACTGTTG

37

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

GCGGATCCAC CATGGGGAG TCGGCC

26

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

AACACCGGGC CGGCGATGAT

20

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Gly Xaa Gly Xaa Xaa Gly
1 5

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

Asp Phe Lys Ser Arg Asn
1 5

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

Asp Leu Lys Ser Lys Asn
1 5

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

Gly Thr Lys Arg Tyr Met
1 5

CLAIMS

1. An isolated protein having a serine/threonine kinase domain, a DFKSRN or DLKSKN sequence in subdomain VIB and/or a GTKRYM sequence in subdomain VIII.
- 5 2. A protein according to claim 1, which additionally comprises an ATP-binding sequence that is Gly-Xaa-Gly-Xaa-Xaa-Gly in subdomain I, and a Lys residue in subdomain II.
3. An isolated protein having a serine/threonine kinase domain which has more than 50% identity to the kinase
10 domain of any of the amino-acid sequences identified herein as SEQ ID. Nos. 2, 4, 6, 8, 10, 12, 14, 16 and 18.
4. A protein according to claim 3, wherein the identity is more than 60%.
5. A protein according to any preceding claim, having
15 serine/threonine kinase activity.
6. An isolated protein having all or part of any of the amino-acid sequences identified herein as SEQ. ID Nos. 2, 4, 6, 8, 10, 12, 14, 16 and 18, and activin receptor type I functionality.
- 20 7. An isolated protein having an amino-acid sequence corresponding to part or all of the amino-acid sequence of an activin type I receptor, and wherein the protein has at least one of the following characteristics:-
 - (i) serine/threonine kinase activity;
 - 25 (ii) activin-binding activity; and
 - (iii) activin type II receptor interaction.
8. An isolated protein having all or part of any of the amino-acid sequences identified herein as SEQ. ID Nos. 2, 4, 6, 8, 10, 12, 14, 16 and 18, and TGF- β -type I receptor
30 functionality.
9. An isolated protein having an amino-acid sequence corresponding to part or all of the amino-acid sequence of a TGF- β -type I receptor, and wherein the protein has at least one of the following characteristics:
 - 35 (i) serine/threonine kinase activity;
 - (ii) TGF- β -binding activity; and
 - (iii) TGF- β -type II receptor interaction.

10. A protein according to any of claims 1 to 5, having all or part of the amino-acid sequence identified herein as SEQ ID No. 2.
11. A protein according to any of claims 1 to 7, having
5 all or part of the amino-acid sequence identified herein as SEQ ID No. 4.
12. A protein according to any of claims 1 to 5, having serine/threonine kinase activity and all or part of the amino-acid sequence identified herein as SEQ ID No. 6.
- 10 13. A protein according to any of claims 1 to 7, having all or part of the amino-acid sequence identified herein as SEQ ID No. 8.
14. A protein according to any of claims 1 to 5, 8 and 9, having all or part of the amino-acid sequence identified
15 herein as SEQ ID No. 10.
15. A protein according to any of claims 1 to 5, having all or part of the amino-acid sequence identified herein as SEQ ID No. 12.
16. A protein according to any of claims 1 to 5, having
20 all or part of the amino-acid sequence identified herein as SEQ ID No. 14.
17. A protein according to any of claims 1 to 7, having all or part of the amino-acid sequence identified herein as SEQ ID No. 16.
- 25 18. A protein according to any of claims 1 to 5, having all or part of the amino-acid sequence identified herein as SEQ ID No. 18.
19. A protein according to any preceding claim, that is a soluble receptor.
- 30 20. An antibody which binds specifically to a protein as defined in any of claims 1 to 19 and not to at least one other such protein.
21. An isolated nucleic acid molecule which codes for, or is complementary to a nucleic acid molecule which codes
35 for, a protein as defined in any of claims 1 to 19.
22. A recombinant nucleic acid molecule comprising at least two heterologous sequences, one of which codes for,

or is complementary to a nucleic acid molecule which codes for, a protein as defined in any of claims 1 to 19.

23. A molecule according to claim 21 or claim 22, wherein the protein is a TGF- β -type I receptor.
- 5 24. A molecule according to claim 21 or claim 22, wherein the protein is an activin receptor.
25. A DNA or RNA/mRNA molecule according to any of claims 21 to 24.
26. A molecule according to any of claims 20 to 24, which
10 additionally comprises, operably associated with the coding sequence, a sequence adapted to allow expression of the protein.
27. A host comprising a molecule according to claim 26, which is capable of expressing the protein.
- 15 28. A host according to claim 27, which comprises PAE cells.
29. A host according to claim 27 or claim 28, transfected with the Chim A receptor plasmid.
30. A product according to any preceding claim, for
20 therapeutic or diagnostic use.
31. Use of a product according to any of claims 1 to 29, for the manufacture of a medicament for use in treating a condition associated with TGF activity.

cons.aa		<u> </u>	<u> </u>	
		G G G V	A K	E
hTGfBR-II		LDTLVGKGRFAEVYKAKLKQNTSEQFETVAVKIFPYDHYASWDRKDI	FSDINLKHENILQF	
mActR-IIB		LLEIKARGRFGCVWKAQLMN-----	DFVAVKIKPLQDKQSWQSEREIFSTPGMKHENLLQF	
mActR-II		LLEVKARGRFGCVWKAQLLN-----	EYVAVKIFPIQDKQSWQNEYEVYSIPGMKHENILQF	
daf-1		LKRVGSGRFGNVSRGDYRG-----	EAVAVKVFNAIDEPAPFKEIEIFETRMLRHPNVLRY	
subdomains		I	II	III IV

hTGfBR-II		LTAERKTELCKQYWLITAFHAKGNLQEYLTRHVISWEDLRNVGSSLARGLSHLSDHTP-C
mActR-IIB		IAAEKRGSNLEVELWLI TAFHDKGSLIDYLGNIITWNE LCHVAETMSRGISYLHEDVPWCR
mActR-II		IGA EKRGTSVDVDLWLI TAFHEKGSLSDFLKANVVS WNELCHIAETMARGLAYLHEDI PGLK
daf-1		IGSDRVDTGFTLWLVIEYHPSGSLHDFLENTVNIETYYNLMRSTASGLAFLHNQIGGSK
subdomains		V VI-A

cons.aa		<u> </u>	<u> </u>	<u> </u>
		DLK N	DFG	
hTGfBR-II		-GRPKMPIVHRDLKSSNILVKNDLTCCCLDFGLSLRL---	GPYSSVDDLANSQVGTARYMAP	
mActR-IIB		GEGHKPSIAHRDFKSKNVLLKSDLTAVLADFG LAVRF---	EPGKPPGD--THGQVGTTRYMAP	
mActR-II		-DGHKPAISHRDIKSKNVLLKNNLTACIADFG LALKF---	EAGKSAGD--THGQVGTTRYMAP	
daf-1		-ESNKPAMAHRDIKSKNIMYKNDLTCAIGDLGLSLSKPEDAASDI IAN--	ENYKCGTVRYLAP	
subdomains		VI-B	VII	VIII

Fig. 1

2/11

a.a C C E G N M C
 5' GCGGATCCTGTTGTGAAGGNAATATGTG 3' Fig. 2A
 BAMHI C C G C

a.a V A V K I F
 5' GCGGATCCGTCGCAGTCAAATTTT 3' Fig. 2B
 BamHI G C G G C
 T T T A

a.a R D I K S K N
 5' GCGGATCCGCGATATTTAAAGCAA 3' Fig. 2C
 BAMHI A C C GTCT
 G A

a.a E P A M Y
 5' CGGAATTCTGGTGCCATATA Fig. 2D
 EcoRI G G G
 A A

M G A A A K L A F A V F L I S C S S G A I L G R A C T R - I I
 M T A P W A L A L A L M O S C A G S C R G E A C C R - I I B
 M E A A V A A P R K G L L L L V L A A A T B R - I I
 M V D G S M I L P V L I M I A L M A L V A L K - 1
 M S M L H G T G M K S O S O Q K K S E A L K - 2
 M A E S A G A S S F P L V V L L A L K - 3
 M L L R S S G K L N V G T K K E A L K - 4
 M G A A A K L A F A V F L I S C S S G A I L G R A C T R - I I
 M T A P W A L A L A L M O S C A G S C R G E A C C R - I I B
 M E A A V A A P R K G L L L L V L A A A T B R - I I
 M V D G S M I L P V L I M I A L M A L V A L K - 1
 M S M L H G T G M K S O S O Q K K S E A L K - 2
 M A E S A G A S S F P L V V L L A L K - 3
 M L L R S S G K L N V G T K K E A L K - 4
 M G A A A K L A F A V F L I S C S S G A I L G R A C T R - I I
 M T A P W A L A L A L M O S C A G S C R G E A C C R - I I B
 M E A A V A A P R K G L L L L V L A A A T B R - I I
 M V D G S M I L P V L I M I A L M A L V A L K - 1
 M S M L H G T G M K S O S O Q K K S E A L K - 2
 M A E S A G A S S F P L V V L L A L K - 3
 M L L R S S G K L N V G T K K E A L K - 4

S E T Q E C L F F Y N A M W E K D R T N Q T G V E P C Y G D K - - D K R R H - C F A I T W K M A C T R - I I
 A E T R E C I Y N A M W E L E R T N Q S G L E R C E G E Q - - D K R R L H - C Y A S W R R N A C C R - I I B
 K F P Q L C K F C D V R A T A L V T C C E S P H C C G N E D - - E K P Q E V - C V A V W R K T B R - I I
 A A A A A L L P S R G P L V M C C E S P H C C G N E D - - C T V T G A - W - Q - C F S S L S I A L K - 2
 T Q G D E K P K V M P K L Y M C C Y C C T S - - H C I T M G G A - G H - C F A I T W K M A L K - 1
 S M E D E L A P E D T L P F L L R C C A C T S - - H C I T M G G A - G H - C F A I T W K M A L K - 2
 A G S G G S G P R G V Q A L L R C C A C T S - - H C I T M G G A - G H - C F A I T W K M A L K - 3
 D G E S T A P T P R P K I L R C C A C T S - - H C I T M G G A - G H - C F A I T W K M A L K - 4
 S E T Q E C L F F Y N A M W E K D R T N Q T G V E P C Y G D K - - D K R R H - C F A I T W K M A C T R - I I
 A E T R E C I Y N A M W E L E R T N Q S G L E R C E G E Q - - D K R R L H - C Y A S W R R N A C C R - I I B
 K F P Q L C K F C D V R A T A L V T C C E S P H C C G N E D - - E K P Q E V - C V A V W R K T B R - I I
 A A A A A L L P S R G P L V M C C E S P H C C G N E D - - C T V T G A - W - Q - C F S S L S I A L K - 2
 T Q G D E K P K V M P K L Y M C C Y C C T S - - H C I T M G G A - G H - C F A I T W K M A L K - 1
 S M E D E L A P E D T L P F L L R C C A C T S - - H C I T M G G A - G H - C F A I T W K M A L K - 2
 A G S G G S G P R G V Q A L L R C C A C T S - - H C I T M G G A - G H - C F A I T W K M A L K - 3
 D G E S T A P T P R P K I L R C C A C T S - - H C I T M G G A - G H - C F A I T W K M A L K - 4

I S S G T I E I V K K Q G C M L D D I M C Y D - - R T D C V - - E K K O S P E V Y F A C T R - I I
 N D E M K X P Y M D F I L P R E - - L H R E Q G K M T C C K A Q - - L R R S I E A L K - 6
 E E G R H P V Y - Q K G C C - F Q - - M K Y - - M P K V E - - L G L - - L R R S I E A L K - 5
 M D Q G G E T T L A S G C C - M K Y - - M P K V E - - L G L - - L R R S I E A L K - 4
 D D Q G G E T T L A S G C C - M K Y - - M P K V E - - L G L - - L R R S I E A L K - 3
 L D S G M P V T S G C C - I G L - - L G L - - L R R S I E A L K - 2
 I S S G T I E I V K K Q G C M L D D I M C Y D - - R T D C V - - E K K O S P E V Y F A C T R - I I
 N D E M K X P Y M D F I L P R E - - L H R E Q G K M T C C K A Q - - L R R S I E A L K - 6
 E E G R H P V Y - Q K G C C - F Q - - M K Y - - M P K V E - - L G L - - L R R S I E A L K - 5
 M D Q G G E T T L A S G C C - M K Y - - M P K V E - - L G L - - L R R S I E A L K - 4
 D D Q G G E T T L A S G C C - M K Y - - M P K V E - - L G L - - L R R S I E A L K - 3
 L D S G M P V T S G C C - I G L - - L G L - - L R R S I E A L K - 2

Fig. 3

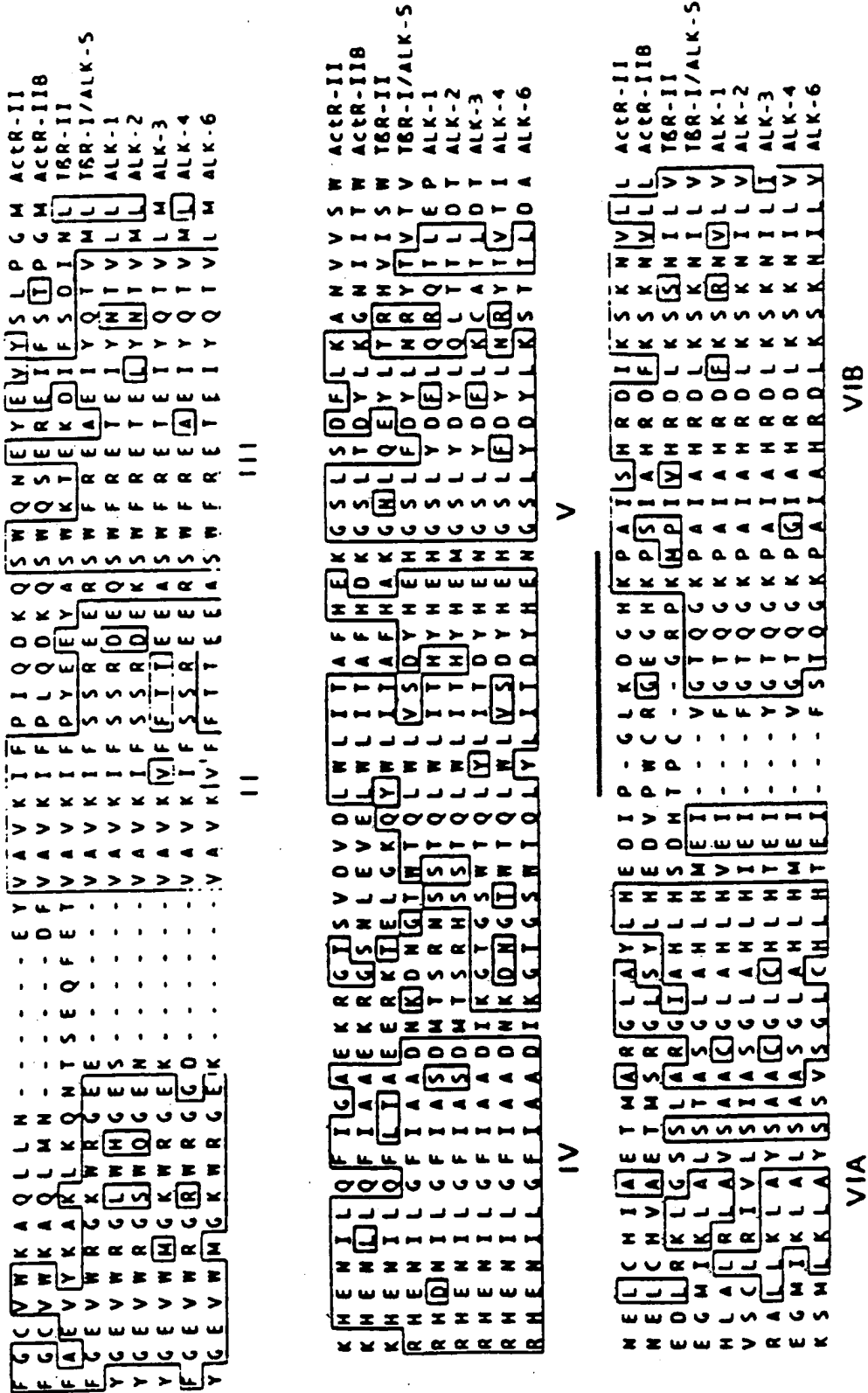


Fig. 3 contd.

6/11

```

K M N L T A C I A D F G L A L L X F E A G K S A G G D D I T H G Q V V G T R R Y M A P E V L E G A C T R - II
K S D L T A Y L A D F G L A A V R R L F E P T L S P V D D L A N S G Q V V G T A R R Y M A P E V L E G A C T R - IIB
K K N G T C C C I A D L G L A V R H M S Q S G S T I D Y L D D I A P N M P R V G T K R R Y M A P E V L D D T B R - II
K S N L Q C C C I A D L G L A V R H M S Q S G S T I D Y L D D I A P N M P R V G T K R R Y M A P E V L D D T B R - I/ALK-5
K K N G Q C C C I A D L G L A V R H M S Q S G S T I D Y L D D I A P N M P R V G T K R R Y M A P E V L D D E A L K - 1
K K N G S C C C I A D L G L A V R H M S Q S G S T I D Y L D D I A P N M P R V G T K R R Y M A P E V L D D E A L K - 2
K K N G M C C C I A D L G L A V R H M S Q S G S T I D Y L D D I A P N M P R V G T K R R Y M A P E V L D D E A L K - 3
K K N G I C C C I A D L G L A V R H M S Q S G S T I D Y L D D I A P N M P R V G T K R R Y M A P E V L D D E A L K - 4
K K N G T C C C I A D L G L A V R H M S Q S G S T I D Y L D D I A P N M P R V G T K R R Y M A P E V L D D E A L K - 6
    
```

VIII

```

A I M F Q R - D A F L R I D M Y A M G L V L W E L A S R C T A A D G P P V D E Y M L P F E E A C C R - II
A I M F Q R - D A F L R I D M Y A M G L V L W E L V S R C K A A D G P P V D E Y M L P F E E A C C R - IIB
R M N L E N A E S F K Q T D V Y S M A L V L W E I A R R R C S I - C G I H E D Y Q L P P F Y D T B R - II
S I M K H F E S E K R A D I I Y A M G L V L W E I A R R R C S I - C G I H E D Y Q L P P F Y D T B R - I/ALK-5
Q I R T D C F E S Y K R V D I I W A F G L V L W E I A R R R C S I - N G I V E D Y R P P F Y D A L K - 1
S L M K N H F O P Y I M A O I Y S F G L I I W E M A R R C I T - G G I V E E Y Q L P P Y N A L K - 3
T I M K H F D S E K C A D I Y A L G L V I W E I A R R C N S - G G V M E E Y Q L P P Y D A L K - 4
S L M R N H F Q S Y I M A D M Y S F G L I I W E I A R R C V S - G G I V E E Y Q L P P Y M D A L K - 6
    
```

X

```

E I G Q H P S L E D M Q E V V V H K K R R P V L R D Y W Q K H A G M L C C E T I E E C W A C C R - II
E I G Q H P S L E E L Q E V V V H K K M R P P T T K D H W L K H P Q G I A A Q L L C C V T I E E C W A C C R - IIB
K V R E H P C S V E S M K R K V D M V L R D R G R P P E I P S F W L N H Q Q I Q M V C E I L T E C W T B R - II
L V P P S D P S S V E E M R K V V C V C E Q K L R P P M T I P N R R L A A D P V L S G L R V M A K I M R E C W T B R - I/ALK-5
V V P P S D P S S F E O M K K V V C V C V D Q Q T P P M I P N R R L A A D P V L S G L R V M A K I M R E C W A L K - 1
V V P S D P S S Y E D M R E K V V V C V C V D Q Q R L R P P I V S N M R W M Q S D E I L R A V L G K L M S E C W A L K - 2
L V P S D P S S I E D M R E K V V V C D Q K L R P P S F P N R M S S D E C I L R Q M G K L M M R T E C W A L K - 3
L V P S D P S S I E D M R E I V C M K K L R R P S S D E C I L R Q M G K L M M R T E C W A L K - 4
L V P S D P S S I E D M R E I V C M K K L R R P S S D E C I L R Q M G K L M M R T E C W A L K - 6
    
```

IX

Fig. 3 contd.

D M D A E A R L S A G C V G E R I T Q M Q R L T M I I T T E D I V T V V T M V T M V D F P A C T R - I I
 D M D A E A R L S A G C V E E R V S L I R R S L V N G T T S D C L V S L V T S V T M V D L L A C T R - I I B
 D M D P E A R L T A Q C V A E R F S E L E H L O R L S G R S C S E E K I P E D G S L N T T T B R - I I
 Y A M G A A R L T A L R I K K T L S Q L S Q P E K P K V I Q (503)
 Y Q N P S A R L T A L R I K K T L L Q K I S M S L D V K I I (532)
 A H N P A S A R L T A L R I K K T L S Q L S Q E D V K I I (505)
 A Q N P A S R L T A L R I K K T L A K M S E S Q Q E D I K L (502)

XI

P K E S S L (513) A C T R - I I
 P K E S S I (536) A C T R - I I B
 K (567) T A R - I I

Fig. 3 contd.

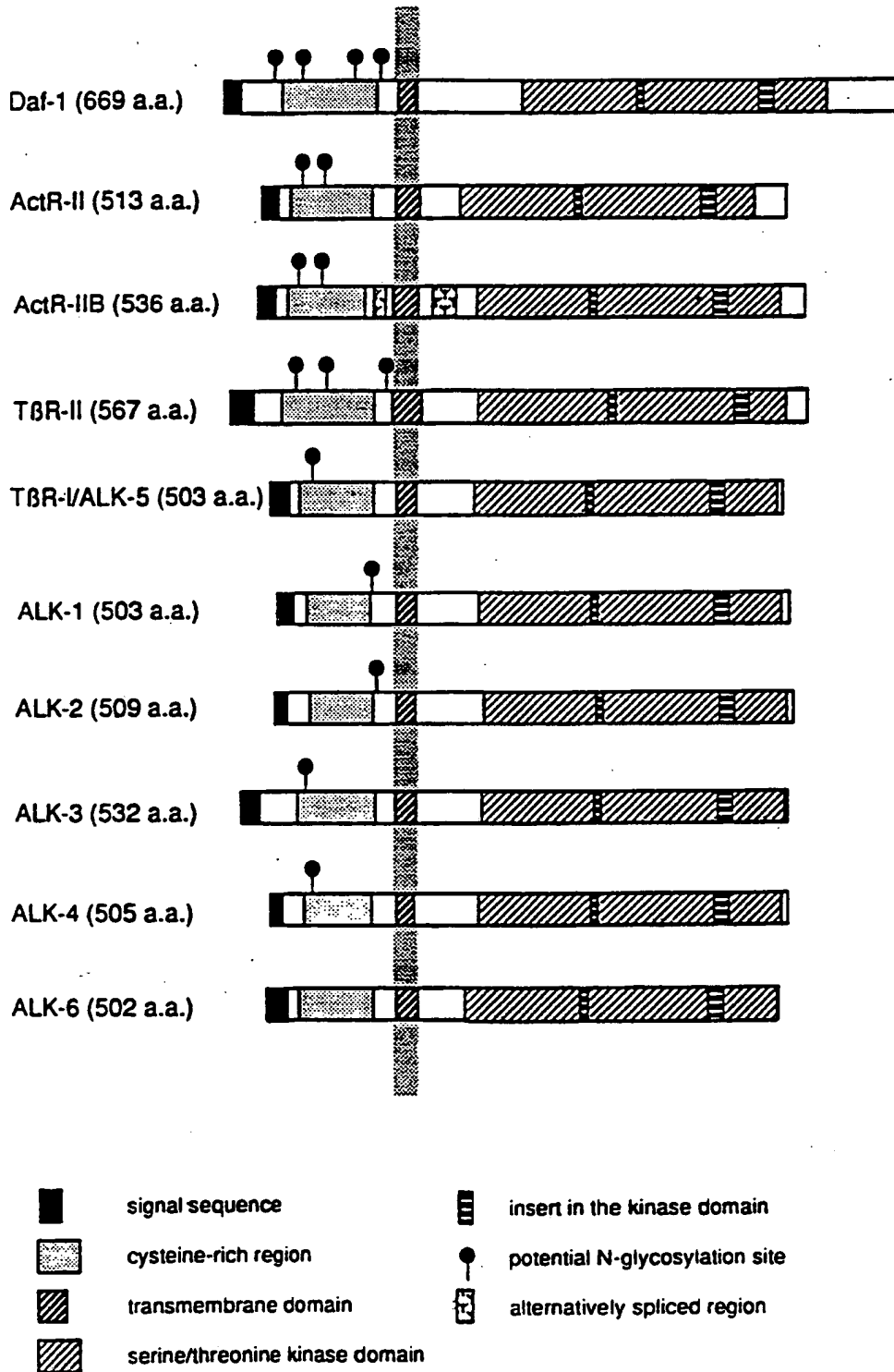


Fig. 4

C - F - - - - C - - - - C X G D - D I - - T C E T - - G - C F V S L - - S D G Majority
 C - T - - - - C E S P - - - - H C K G P - S C G N E D H C E G Q R - G A W C T V V L V R R E E G
 C - Y - - - - S G - - - - P D Q A - A Y - - T C E I T T - G Q - Q N G A H C F S S L S I N D D G
 C - A - - - - C T S - - - - P D Q A - A Y - - T C E I T T - G Q - Q N G A H C F S S L S I N D D G
 C - F F N - - - - A N W E K D R T N Q T G V E P C E M S N C S I - - K R R H C F A T - - D G L C F V S V T T D
 C - I Y N - - - - A N W E L E R T N Q S G L E R C E M S N C S I - - K R R H C F A T - - D G L C F V S V T T D
 C - K E - - - - D V R F S T - - - - A N W E L E R T N Q S G L E R C E M S N C S I - - K R R H C F A T - - D G L C F V S V T T D
 C - H C S R E V G C - - - - D V R F S T - - - - A N W E L E R T N Q S G L E R C E M S N C S I - - K R R H C F A T - - D G L C F V S V T T D
 I E I V E K G C - - - - C Y D R T L - - - - G S P F - - C V K S P K S P G - T V T E C - C E G D L C Majority
 R H P Q E H R G - - - - C G N L H - - - - R E L C - - R G R P T E - - F V N H Y C C D S H I C
 F H V Y Q K G G - - - - F O V Y E - - - - G K M T C - - K T P P S P G Q A R R T I E C C C Q G D W C
 G E T T L A S G - - - - I P K V E L V P R D R P E - - - - G S D F F Y C C L S S E D - - - - C R T N L C
 M E H H V R T C - - - - I A E I D L I - - - - V P A G K P F F V C A P S S K T G S V T T Y C C C N Q D Y C
 K V I H N S M C - - - - C Y D R T D - - - - C V A E K E N P P O - - - - V Y F F M C C C E G N M F C
 I E I V K Q G G C W L D D D I N C Y D R Q E - - - - C V A E K E N P P O - - - - V Y F F M C C C E G N M F C
 I F L V K K G C W L D D D I N C Y D R Q E - - - - C V A E K E N P P O - - - - V Y F F M C C C E G N M F C
 N I T L E T V C H D P K L P Y H D F I L E E D A A S R K C I M K E K R K P G E T F F W I C C S S D E C
 P E I S H F G C - - - - M D E K S V T D E T E F H D T A A K V C T N N T K D P H A T V C C D K G N F C

Fig. 5

ALK-2	ALK-3	ALK-4	ALK-5	ActR-II	ActR-IIB	TBR-II	daf-1	
79	60	61	63	40	40	37	39	ALK-1
	63	64	65	41	39	37	39	ALK-2
		63	65	41	38	37	39	ALK-3
			90	41	40	39	42	ALK-4
				42	40	41	43	ALK-5
					78	48	35	ActR-II
						47	32	ActR-IIB
							34	TBR-II

Fig. 6

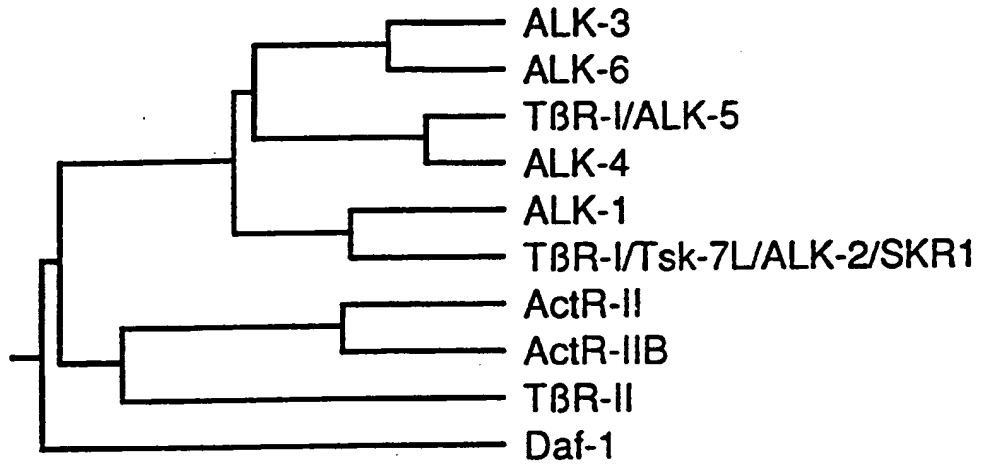


Fig. 7

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT OR DRAWING
- BLURRED OR ILLEGIBLE TEXT OR DRAWING
- SKEWED/SLANTED IMAGES
- COLOR OR BLACK AND WHITE PHOTOGRAPHS
- GRAY SCALE DOCUMENTS
- LINES OR MARKS ON ORIGINAL DOCUMENT
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.